

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 09:02:32 ; Search time 7 seconds  
(without alignments)  
1294.310 Million cell updates/sec

Title: US-09-843-250-2  
Perfect score: 2408  
Sequence: 1 MNYNKILVSEGLSQKHLI.....AEFEHASSTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues  
Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PTCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PTCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	736	30.6	453	10	US-09-815-242-10253
2	376	15.6	490	9	US-09-738-626-6140
3	375.5	15.6	385	10	US-09-815-242-11692
4	340.5	14.1	424	10	US-09-815-242-5097
5	185	7.7	35	10	US-09-776-490-19
6	185	7.7	35	10	US-09-776-491-19
7	174	7.2	35	10	US-09-776-490-20
8	174	7.2	35	10	US-09-776-491-20
9	168	7.0	35	10	US-09-776-490-21
10	168	7.0	35	10	US-09-776-491-21
11	118.5	4.9	354	1	US-08-976-063C-4
12	110	4.6	35	10	US-09-776-490-12
13	110	4.6	35	10	US-09-776-491-12
14	109	4.5	35	10	US-09-776-490-14
15	109	4.5	35	10	US-09-776-491-14
16	108	4.5	35	10	US-09-776-490-13
17	108	4.5	35	10	US-09-776-490-15
18	108	4.5	35	10	US-09-776-491-13
19	108	4.5	35	10	US-09-776-491-15

20	107.5	4.5	548	9	US-10-047-542-78
21	106.5	4.4	951	9	US-09-924-097-15
22	106	4.4	35	10	US-09-776-490-9
23	106	4.4	35	10	US-09-776-491-9
24	104	4.3	35	10	US-09-776-490-18
25	104	4.3	35	10	US-09-776-491-18
26	103	4.3	35	10	US-09-776-490-11
27	103	4.3	35	10	US-09-776-491-11
28	102	4.2	35	10	US-09-776-490-8
29	102	4.2	35	10	US-09-776-490-10
30	102	4.2	35	10	US-09-776-491-8
31	102	4.2	35	10	US-09-776-491-10
32	101.5	4.2	376	9	US-09-738-626-6115
33	101	4.2	17	10	US-09-776-490-45
34	101	4.2	17	10	US-09-776-490-45
35	101	4.2	17	10	US-09-776-491-44
36	101	4.2	17	10	US-09-776-491-45
37	101	4.2	3069	9	US-09-712-363-246
38	100	4.2	35	10	US-09-776-490-17
39	100	4.2	35	10	US-09-776-491-17
40	95	3.9	17	10	US-09-776-490-46
41	95	3.9	17	10	US-09-776-491-46
42	94.5	3.9	282	10	US-09-925-301-1362
43	94	3.9	917	10	US-09-815-242-5603
44	94	3.9	920	10	US-09-815-242-12181
45	94	3.9	920	10	US-09-815-242-12995

ALIGNMENTS

RESULT 1  
US-09-815-242-10253  
; Sequence 10253, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10253  
; LENGTH: 453  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10253

Query Match 30.6%; Score 736; DB 10; Length 453;  
Best Local Similarity 36.4%; Pred. No. 1.1e-57;

Matches	165;	Conservative	64;	Mismatches	174;	Indels	50;	Gaps	8;
Qy	20	IHGDEELFOHELUKTI	IFARNWLFJLTHDSL	ILAPGDYVYAKMGID	DEVIVSRONDGSI	RAFLN	79		
Db	24	IYDTPDYOLELERIF	IGRCWFLFLAHESQ	IKPGDFFNTYMG	EDAVVYVRQKGS	IKAFLN	83		
Qy	80	VCRRHKTIIVSYEAG	NAGFYCSYHGWGFG	SGNSLOSVPEKDL	YGESLNKKCLGL	KEVA	139		
Db	84	QCRHRAMRVSYAD	CCNRAFTCPVHG	SYGINGELIDV	LEPRAYPGGLCK	SHWGLNEVP	143		
Qy	140	RVESFHFYICDQEA	EAPPLMDYLGDAA	WLPEMF-KHSG	GLELVGPGKVY	IKANWKP	198		
Db	144	CVESYKGLIFGN	WDTSAFGLRDY	LIDIAWYLDG	MLDRREGTEI	GVGVQKWVINC	WKP	203	
Qy	199	AENFYGDATHV	GWTHASSL-----	RGSEIFSS	LAGNAALP-----		234		
Db	204	AEQFASDQYHAL	FSHASAVQVLG	AKDGDGSKRLGD-----	GQTARPYMETAK	DALQFG	256		
Qy	235	--PEGAGLQMT	SKYSGGMGLWD	GYSGVHSADL	VPMLAFGGAQ	BERLNKETGDV	VRARY	292	
Db	257	QDGHGSGFFTE	KPDANVWV--	DGAVSSY	REYIAE-----	AQRGLGEVRAL	RL	303	
Qy	293	RSILNCTVFP	NNSMTCVSGV	PKVWNPIDANT	EVWYVAIVEK	MDPDLKRR	LSDVQRTV	352	
Db	304	AGHNN--	IPFTLSWLNG	TATLRV	WHPRGPDQVE	WFAFCITDKA	ASDEVKAAPENS	ATRAF	361
Qy	353	GPAGFWESDD	NNMETASQ	NCKKYQSRD	SDLLSNLGLG	GEDVYDAVY	PGVGVCKSAIGTS	412	
Db	362	GPAGFLEQD	SDSENWCEIQ	KLKGHARN	SKLCEMLG	LQGEKRR	DDGIPGIT--NYIF	SETA	420
Qy	413	YRGFYRAYQAH	VSSSNNAE	FEHASST	WHTBLTK	445			
Db	421	ARGMYORMAD	LSSNESW	OEVLDTK	TAAYQOE	VWVK	453		

```

RESULT 2
US-09-738-626-6140
; Sequence 6140, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6140
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6140

Best Match 15.6%; Score 376; DB 9; Length 490;
Query Local Similarity 26.3%; Pred.No.1.3e-25;
Matches 118; Conservative 66; Mismatches 215; Indels 50; Gaps 15;

QY 20 IHGDELFQHELKLTIFARNWLFTHDSLIPAPGDYVTAKMGIDDEVIVSQNDGSIKRAFLN 79

```

[illegible]

```

RESULT 3
US-09-815-242-11692
; Sequence 11692, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; PROKARYOTES
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11692
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-815-242-11692

```







COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-491-20

Query Match 7.28; Score 174; DB 10; Length 35;  
Best Local Similarity 85.78; Pred. No. 2.7e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCHRGRKTLVSEAGNAKGFVCSYHGWFSGNGE 113  
|||||:|||||:|||||:|||||:|||||:  
Db 1 NVCHRGRKTLVSEAGNAKGFVCSYHGWFSGNCK 35

RESULT 9  
US-09-776-490-21  
Sequence 21, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-490-21

Query Match 7.08; Score 168; DB 10; Length 35;  
Best Local Similarity 80.08; Pred. No. 9.3e-09;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCHRGRKTLVSEAGNAKGFVCSYHGWFSGNGE 113  
|||||:|||||:|||||:|||||:|||||:  
Db 1 NVCHRGRKTLVSEAGNAKGFVCSYHGWFSGNCK 35

RESULT 10  
US-09-776-491-21  
Sequence 21, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-491-21

Query Match 7.08; Score 168; DB 10; Length 35;  
Best Local Similarity 80.08; Pred. No. 9.3e-09;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;



```
;
; CITY: Charlotte
; STATE: NO. US20010013135A1th Carolina
; COUNTRY: USA
; ZIP: 28234
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/776,491
; FILING DATE: 02-Feb-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/810,009
; FILING DATE: 04-MAR-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-776-491-12

Query Match 4.6%; Score 110; DB 10; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.0013;
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRHGRKTLVSVNAGNAKGFVCSYHGWGFGSNG 112
| | | | | : | | | | | | | | | | : |
Db 1 NQCRHGRMCRADGNAGNAKAPTCSYHGWAYDTAG 34

RESULT 14
US-09-776-490-14
; Sequence 14, Application US/09776490
; Patent No. US20010012886A1
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
;          Johan, Gurmukh S.
;          Gray, John
;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
;
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: NO. US20010012886A1th Carolina
; COUNTRY: USA
; ZIP: 28234
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/776,490
; FILING DATE: 02-Feb-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/810,009
; FILING DATE: 04-MAR-1997
```

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-776-490-14

Query Match 4.5%; Score 109; DB 10; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.0016;
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHGRKTLVSVNAGNAKGFVCSYHGWGFGSNG 112
| | | | | : | | | | | | | | | | : |
Db 1 NQCRHGRMCRADGNAGNAKAPTCSYHGWAYDSAG 34

RESULT 15
US-09-776-491-14
; Sequence 14, Application US/09776491
; Patent No. US20010013135A1
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
;          Johan, Gurmukh S.
;          Gray, John
;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
;
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: NO. US20010013135A1th Carolina
; COUNTRY: USA
; ZIP: 28234
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/776,491
; FILING DATE: 02-Feb-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/810,009
; FILING DATE: 04-MAR-1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
```

US-09-776-491-14

Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred No. 0.0016;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Qy 79 NVCRHRGKTLVSVEAGNAKGFCVCSYHGWFSGNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NQCRHRGMRICRADGGNAKSPCTCSYHGWAYDSAG 34

Search completed: January 27, 2003, 09:20:54  
Job time : 8 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	224	9.3	446	4	US-09-004-393B-4		Sequence 4, Appli
2	223	9.3	439	4	US-09-004-393B-2		Sequence 2, Appli
3	185	7.7	35	4	US-08-810-009-19		Sequence 19, Appl
4	174	7.2	35	4	US-08-810-009-20		Sequence 20, Appl
5	168	7.0	35	4	US-08-810-009-21		Sequence 21, Appl
6	112.5	4.7	379	3	US-09-028-934-36		Sequence 36, Appl
7	110	4.6	35	4	US-08-810-009-12		Sequence 12, Appl
8	109	4.5	35	4	US-08-810-009-14		Sequence 14, Appl
9	109	4.5	622	4	US-09-311-626B-4		Sequence 4, Appli
10	108	4.5	35	4	US-08-810-009-13		Sequence 13, Appl
11	108	4.5	35	4	US-08-810-009-15		Sequence 15, Appl
12	106	4.4	35	4	US-08-810-009-9		Sequence 9, Appli
13	104	4.3	35	4	US-08-810-009-18		Sequence 18, Appl
14	103.5	4.3	432	4	US-08-809-326A-16		Sequence 16, Appl
15	103.5	4.3	649	4	US-08-809-326A-15		Sequence 15, Appl
16	103	4.2	35	4	US-08-810-009-11		Sequence 11, Appl
17	102	4.2	35	4	US-08-810-009-8		Sequence 8, Appli
18	102	4.2	35	4	US-08-810-009-10		Sequence 10, Appl
19	101	4.2	17	4	US-08-810-009-44		Sequence 44, Appl
20	101	4.2	17	4	US-08-810-009-45		Sequence 45, Appl
21	100	4.2	35	4	US-08-810-009-17		Sequence 17, Appl
22	99.5	4.1	256	4	US-09-325-932A-57		Sequence 57, Appl
23	95	3.9	17	4	US-08-810-009-45		Sequence 46, Appl
24	94.5	3.9	350	2	US-08-828-922-1		Sequence 1, Appli
25	94	3.9	560	3	US-08-814-052-6		Sequence 6, Appli
26	94	3.9	560	3	US-08-812-829-6		Sequence 6, Appli
27	94	3.9	563	4	US-09-134-001C-4800		Sequence 4800, Ap

APPLICANT: Rathinasabapathi, Bala  
APPLICANT: Burnet, Michael  
TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and  
TITLE OF INVENTION: Plants Transformed Therewith  
FILE REFERENCE: US-162  
CURRENT APPLICATION NUMBER: US/09/004.393B  
PRIOR FILING DATE: 1998-01-08  
PRIOR APPLICATION NUMBER: 60/035,147  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 439  
TYPE: PRT  
ORGANISM: Spinacia oleracea  
US-09-004-393B-2

Query Match 9.3%; Score 223; DB 4; Length 439;  
Best Local Similarity 30.1%; Pred. No. 5.7e-13;  
Matches 58; Conservative 38; Mismatches 83; Indels 14; Gaps 7;  
QY 27 FOHELKTIARNLWFLTHSLIPACDYVTAKMGIDIVISQNDGSIKRAFLNVCRRHGK 86  
DB 108 YSHELRIPIYKGMQVAGISDQKEPNQYFTSLGNVEXLVSRDGEKGVHAFHNVCTHRA- 166  
QY 87 TLVSVEAGNAKGFVCSYHGWGFGSGNGLQSVPFKDLXGESLNKKCLGLKEVARVESPHG 146  
DB 167 SILACSGKSKCFVPCYHGWVYMGDSLAKASKAP--EQNLDPKELGLVPL-KVAVWGP 223  
QY 147 FYGCFDQAPFLMD----YLGDAAWYLEPMFKHS--GGLELVGPPGVKVIKANWKAPAE 200  
DB 224 FVLISLDRSLEEGDVGTEWLGTSA---EDVKAHAFDPSLQFI-HRSEFPNESNKKIFSD 279  
QY 201 NFVGDAYHYGVNTH 213  
DB 280 NYLDDSYHYVPIAH 292

RESULT 3  
US-08-810-009-19  
Sequence 19, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810.009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-19

Query Match 7.7%; Score 185; DB 4; Length 35;  
Best Local Similarity 94.3%; Pred. No. 5.2e-11;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSGNGL 113  
DB 1 NVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSGNGK 35

RESULT 4  
US-08-810-009-20  
Sequence 20, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810.009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-20

Query Match 7.2%; Score 174; DB 4; Length 35;  
Best Local Similarity 85.7%; Pred. No. 5.8e-10;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSGNGL 113  
DB 1 NVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSGNGK 35

RESULT 5  
US-08-810-009-21

```
; Sequence 21, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-21

Query Match 7.0%; Score 168; DB 4; Length 35;
Best Local Similarity 80.0%; Pred. No. 2.2e-09;
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWGFGSNGE 113
Db 1 NVCRHRGKTVDAEAGNAKGPVCGYHGWYGSNGK 35

RESULT 6
US-09-028-934-36
; Sequence 36, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrrolitritin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-028-934-36

Query Match 4.7%; Score 112.5; DB 3; Length 379;
Best Local Similarity 20.3%; Pred. No. 0.017;
Matches 95; Conservative 54; Mismatches 149; Indels 171; Gaps 26;

QY 20 IHGDEELFQ-----HELTIFARNW-LFLTHDSLIPAGDYVTAKMGIDVIVSRQDGS 73
Db 5 IHQEPERIROAGSVNDLTTQTAAASWYVAMRSDALRGKPK---VAIKLFGQPLVAWRDGGGR 61
QY 74 IRAFLNVCRHRGKTLVSEAGNAKGFV-----CSYHGWGFGSNGELQSVFFPKDLYGES 127
Db 62 PVVMERYCSHLGASL-----AKGVVEGCIQCPFNHRYDSTGACSHVP-----GHS 108
QY 128 LNKKCLGKLEVARVESF-----HGFIYCCFQDEAP-----PLMDYLGDAAWY 169
Db 109 -----TEVPLEPIPTPTAROSVYPMERYGVVWYGTAKPLPLPMPAESSESH 160
QY 170 LEPMFHSGGLELVGPPGKVVIKANNKAPAEVPGDAYHVGVTHASLSRGE-SIF---S 225
Db 161 QSLRFAYETTTSVL---RII-----ENFY-DAQHAAPVHQLPISAFELKLFDESS 206
QY 226 SLAGNAALPPE---GAGLQM-TSKYSGMGVL-----WDGYSGVHSDLV 266
Db 207 PPPQEQALARDGAWFGAGIDFHVDRYFGLGVISRTLGLSMSRMQLHFDGYPG----- 259
QY 267 PELMAFGGAKQERLKEIGDVRARIYRSHLNCVFPNNSMLTCSGVFKVMNPIDANTTEV 326
Db 260 -----GCIMTVSLD---GDVKYRL---LQCV-----TPVDKKEETVM 289
QY 327 WTYAIVEKDMPEDLKRRLDVSV---QRTVGPAGF---WESDDNDNMETASQNGKKYQSR 379
Db 290 HMLLAIKKG--DGVVRSAAFNILYGLQTWAAAGYDVAINMSKAD-----GGGAFSKY 340
QY 380 DSDLLSNLGFGEVDYGDVYGVVGVGKSAIGETSYRGFYRAYQAHVSSN 428
Db 341 DQLIL-----KYRAFYRRWNVKNVALEN 362

RESULT 7
US-08-810-009-12
; Sequence 12, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
```

```
/
/ APPLICANT: Johal, Gurmukh S.
/ APPLICANT: Gray, John
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
/ TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
/ NUMBER OF SEQUENCES: 65
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: BELL, SELTZER, PARK & GIBSON
/ STREET: P.O. Drawer 34009
/ CITY: Charlotte
/ STATE: No. 6211437th Carolina
/ COUNTRY: USA
/ ZIP: 28234
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/810,009
/ FILING DATE: 04-MAR-1997
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Spruill, W. Murray
/ REGISTRATION NUMBER: 32,943
/ REFERENCE/DOCKET NUMBER: 5718-4
/ TELEPHONE: 919-881-3140
/ TELEFAX: 919-881-3175
/ TELEX: 575102
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 35 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-810-009-12

Query Match 4.6%; Score 110; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.00077;
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRRHGTLSVEAGNAKGFVCSYHGWFSGNSG 112
| | | | | : | | | | | | | | | | : |
Db 1 NQCRHRGMRICRADGNAKAPTCSYHGWAYDTAG 34

RESULT 8
US-08-810-009-14
/ Sequence 14, Application US/08810009
/ Patent No. 6211437
/ GENERAL INFORMATION:
/ APPLICANT: Briggs, Steven P.
/ APPLICANT: Johal, Gurmukh S.
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
/ TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
/ NUMBER OF SEQUENCES: 65
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: BELL, SELTZER, PARK & GIBSON
/ STREET: P.O. Drawer 34009
/ CITY: Charlotte
/ STATE: No. 6211437th Carolina
/ COUNTRY: USA
/ ZIP: 28234
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/810,009
/ FILING DATE: 04-MAR-1997
```

```
/
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Spruill, W. Murray
/ REGISTRATION NUMBER: 32,943
/ REFERENCE/DOCKET NUMBER: 5718-4
/ TELEPHONE: 919-881-3140
/ TELEFAX: 919-881-3175
/ TELEX: 575102
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 35 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-810-009-14

Query Match 4.5%; Score 109; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.00096;
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRRHGTLSVEAGNAKGFVCSYHGWFSGNSG 112
| | | | | : | | | | | | | | | | : |
Db 1 NQCRHRGMRICRADGNAKAPTCSYHGWAYDSAG 34

RESULT 9
US-09-311-626B-4
/ Sequence 4, Application US/09311626B
/ Patent No. 6399347
/ GENERAL INFORMATION:
/ APPLICANT: Jorgensen, Per Lina
/ APPLICANT: Schnorr, Kirk
/ APPLICANT: Andersen, Lene No. 6399347boe
/ APPLICANT: Schulein, Martin
/ APPLICANT: Outtrup, Helle
/ TITLE OF INVENTION: No. 6399347el Rhamnogalacturonan Hydrolases
/ FILE REFERENCE: 5572.204-US
/ CURRENT APPLICATION NUMBER: US/09/311,626B
/ CURRENT FILING DATE: 1999-05-13
/ PRIOR APPLICATION NUMBER: 0608/98
/ PRIOR FILING DATE: 1998-05-01
/ PRIOR APPLICATION NUMBER: 60/084,358
/ PRIOR FILING DATE: 1998-05-05
/ NUMBER OF SEQ ID NOS: 61
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 622
/ TYPE: PRT
/ ORGANISM: Bacillus licheniformis
/ US-09-311-626B-4

Query Match 4.5%; Score 109; DB 4; Length 622;
Best Local Similarity 21.7%; Pred. No. 0.078;
Matches 92; Conservative 50; Mismatches 130; Indels 152; Gaps 23;

QY 84 RGKTLVSVEAG--NAKGFVCS-----YHGWFSGNSELQSVFPEK-----DLYG 125
: | | : | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 245 KGKVGIGNADYRNAQGRILSGPEYLTVPFKG---DTCAELTTVNYEPARGNVADMGDSYG 301
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 126 ESLNKKCLGL-----KEVARVESPHGFIYG-----CFDOEAPPLMD 161
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 302 NRVDRELAGVAYLDGERPSFVMARGYVTRVLVAYNFRG---GKLTKLWTFDSAPNGA 358
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 162 YLGDAAWYLEPMFKHSGGLVLPVPPGKVIKANWKAPEAFVGDYHVWTHASLSRGE 221
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 359 YAGQG-----NHSLSVADVDDGDKDEIIYGAMAVDHDGKG--LYSTGWGHGDAHMTGN 409
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 222 SISSLAGNAALPEGAGLQMTSKYSGCMGLWMDGYSGVHSADLVPELMFAGGAKQERL- 280
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 410 -----LDPSRPGLEVFQ-----VHENSNSPYGLSPRDARTGKII 443
: | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```



QY	281	-----NKEIG-DYRARIYRSHLNCITVFPNNSMITCSGVFKVNPIDANTT-EWYTYAIVE	333
Db	444	WGVIHAGDVGVRGMAADIDPRYEGAEYWAWSLYTAKGV-KIGNTLPSSTNFGII	496
QY	334	KDMPEDLKRLADSVQ-----RTVCPAGFWESDDDNMETAS---QNG-KKYQSR	379
Db	497	--WDGDLQRELLDSNRIDKWDYQNSRTV-----NLTASGASANNGTATPSL	542
QY	380	DSLLSN-----LGFGEED-----VYGDVYPCWGVKSAIG--ETSY	413
Db	543	QADILGDWREEVVRADSSSELRIYTTTDTVTRMYTLMHDVAYRLGIAWQVNGVNPQH	602
QY	414	RGFY 417	
Db	603	TGFY 606	

```

RESULT 10
US-08-810-009-13
; Sequence 13, Application us/08810009
; Patent No. 5211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS

```

```

/ TELEX: 575102
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 35 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-810-009-13

Query Match 4.5%; Score 108; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.0012;
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSVENAGNAGFCVCSYHGWFQSGNG 112
| | | | | : | | | | | | | | | : |
Db 1 NQCRHRGMRICRSDAGNAKAPTCSYHGWAYDTAG 34

```

```

; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; NUMBER OF SEQUENCES: 65
; NUMBER OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: NO. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprull, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-810-009-15

Query Match      4.5%; Score 108; DB 4; Length 35;
Best Local Similarity 50.0%; Pred. No. 0.0012;
Matches 17; Conservative 4; Mismatches 13; Indels 0;

QY          79   NVCRHRGKTLVSVEAGNAKGFCVSYHGWFSGNS 112
Db          1    NQCRRHGRMCIRADGGNKAASPTCSYHGWAYDTGG 34

```

```

RESULT 12
US-08-810-009-9
; Sequence 9, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/810.009  
;; FILING DATE: 04-MAR-1997  
;; CLASSIFICATION: 800  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sprull, W. Murray  
;; REGISTRATION NUMBER: 32,943  
;; REFERENCE/DOCKET NUMBER: 5718-4  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 919-881-3140  
;; TELEFAX: 919-881-3175  
;; TELEX: 575102  
;; INFORMATION FOR SEQ ID NO: 9:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 35 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-810-009-9

Query Match 4.4%; Score 106; DB 4; Length 35;  
Best Local Similarity 51.4%; Pred. No. 0.0019;  
Matches 18; Conservative 4; Mismatches 13; Indels 0; Gaps 0;  
Qy 79 NVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSGNE 113  
| | | | | : | | | | | | | | | : | :  
Db 1 NQCRHGMICRSDAGNAKAPTCVHGWAYDIAGK 35  
;;  
;; RESULT 13  
US-08-810-009-18  
;; Sequence 18, Application US/08810009  
;; Patent No. 6211437  
;; GENERAL INFORMATION:  
;; APPLICANT: Briggs, Steven P.  
;; APPLICANT: Johal, Gurmukh S.  
;; APPLICANT: Gray, John  
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
;; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
;; NUMBER OF SEQUENCES: 65  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESS: BELL, SELTZER, PARK & GIBSON  
;; STREET: P.O. Drawer 34009  
;; CITY: Charlotte  
;; STATE: No. 6211437th Carolina  
;; COUNTRY: USA  
;; ZIP: 28234  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/810.009  
;; FILING DATE: 04-MAR-1997  
;; CLASSIFICATION: 800  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Sprull, W. Murray  
;; REGISTRATION NUMBER: 32,943  
;; REFERENCE/DOCKET NUMBER: 5718-4  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 919-881-3140  
;; TELEFAX: 919-881-3175  
;; TELEX: 575102  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 35 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-810-009-18

Query Match 4.3%; Score 104; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0029;  
Matches 18; Conservative 2; Mismatches 14; Indels 0; Gaps 0;  
Qy 79 NVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSGNG 112  
| | | | | : | | | | | | | | | : | :  
Db 1 NSCRHRGALLCPFSKGNQKHFVCRYHGWSYDSSG 34

;;  
;; RESULT 14  
US-08-809-326A-16  
;; Sequence 16, Application US/08809326A  
;; Patent No. 6165478  
;; GENERAL INFORMATION:  
;; APPLICANT: Izutsu, Hiroshi  
;; APPLICANT: Obara, Kazuhiko  
;; APPLICANT: Matsumoto, Akira  
;; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS, TRANSFORMA  
;; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS, TRANSFORMA  
;; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA  
;; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
;; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM  
;; TITLE OF INVENTION: OF ANTI-BODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
;; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
;; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
;; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
;; TITLE OF INVENTION: PNEUMONIAE GENE  
;; NUMBER OF SEQUENCES: 31  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESS: Pennie & Edmonds LLP  
;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: USA  
;; ZIP: 10036-2711  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/809.326A  
;; FILING DATE: 19-MAR-1997  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 224711/94  
;; FILING DATE: 20-SEP-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 106006/95  
;; FILING DATE: 28-APR-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 106008/95  
;; FILING DATE: 28-APR-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 106009/95  
;; FILING DATE: 28-APR-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 106010/95  
;; FILING DATE: 28-APR-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 106011/95  
;; FILING DATE: 28-APR-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Miller, Charles E.  
;; REGISTRATION NUMBER: 24,576  
;; REFERENCE/DOCKET NUMBER: 7426-043-999  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 790-9090  
;; TELEFAX: (212) 869-8864/9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 16:  
;; SEQUENCE CHARACTERISTICS:

; PRIOR APPLICATION DATA:

Job time : 12.166 / secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:54:32 ; Search time 25.8333 Seconds  
(without alignments)  
3581.232 Million cell updates/sec

Title: US-09-843-250-2  
Perfect score: 2408  
Sequence: 1 MNYNKKILVSEGLSQKHLI.....AEFEHASTWHTELTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*  
1: sp-archaea.\*  
2: sp-bacteria.\*  
3: sp-fungi.\*  
4: sp-human.\*  
5: sp-invertebrate.\*  
6: sp-mammal.\*  
7: sp-mhc.\*  
8: sp-organelle.\*  
9: sp-phage.\*  
10: sp-plant.\*  
11: sp-prodent.\*  
12: sp-virus.\*  
13: sp-vertebrate.\*  
14: sp-unclassified.\*  
15: sp-rvirus.\*  
16: sp-bacteriap.\*  
17: sp-archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2331	96.8	449	2 Q938R9	Q938r9 pseudomonas
2	2291	95.1	449	2 Q52142	Q52142 pseudomonas
3	2207	91.7	449	2 Q92173	Q92173 pseudomonas
4	2192	91.0	447	2 Q915D2	Q915d2 comamonas t
5	2188	90.9	447	2 O52382	O52382 ralstonia s
6	2075	86.2	447	2 O8VUD4	O8vud4 burkholderi
7	2055	85.3	447	2 P95564	P95564 pseudomonas
8	2026	84.1	447	2 Q8RTL4	Q8rtl4 comamonas s
9	1968.5	81.7	451	2 Q45695	Q45695 burkholderi
10	1490	61.9	277	2 Q9ETK2	Q9etk2 pseudomonas
11	1488	61.8	277	2 Q9F5S4	Q9f5s4 pseudomonas
12	1486	61.7	277	2 Q9F5T3	Q9f5t3 pseudomonas
13	1484	61.6	277	2 Q9F5S8	Q9f5s8 pseudomonas
14	1483	61.6	277	2 Q9F5S9	Q9f5s9 pseudomonas
15	1451	60.3	277	2 Q9F5S3	Q9f5s3 pseudomonas
16	1405	58.3	277	2 Q9EUC7	Q9euc7 pseudomonas

RESULT 1

Q938R9 ID Q938R9 PRELIMINARY; PRT; 449 AA.

AC Q938R9; DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Naphthalene dioxygenase.

GN NAHAC.

OS Pseudomonas fluorescens.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxID=294;

RN [1]

RP SEQUENCE FROM N.A.

RA Min K.-H., Ji S.-H.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY048759; AAL07262.1; -.

DR InterPro: IPR001281; Rieske.

DR InterPro: IPR001663; Ring\_hydroxyl\_A.

DR Pfam: PF00355; Rieske; 1.

DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; UNKNOWN\_1.

DR Dioxigenase.

SW SEQUENCE 449 AA; 49608 MW; 5EFEDBE0282FE812 CRC64;

Query Match 96.8%; Score 2331; DB 2; Length 449;  
Best Local Similarity 96.0%; Pred. No. 3 9e-171;  
Matches 431; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

QY	1	MNYNKKILVSEGLSQKHLIHGDEELFQHELRTIFARNMLFLTHDSLIPAPGDVYTKMG	60
Db	1	MNYNKKILVSEGLTQKHLIHGDEELFQHELRTIFARNMLFLTHDSLIPSGDVYTKMG	60
QY	61	IDEIVSRQNDGSIIRAFNLVNCVRRHRTILVSVEAGNAKGFVCSYHGFGSGNGELQSVPE	120
Db	61	IDEIVSRQSDGSIIRAFNLVNCVRRHRTILVNAEAGNAKGFVCSYHGFGSGNGELQSVPE	120
QY	121	KDLYGESLNKCLGLKEARVESFHGFCFQDQEAAPPLMDYLGDAAWYLEPMPFKHSGGL	180

Q9f5t8 pseudomonas  
Q9f5t5 pseudomonas  
Q9etv3 pseudomonas  
Q9f5s5 pseudomonas  
Q9f5t6 pseudomonas  
Q9f5s7 pseudomonas  
Q9f5t4 pseudomonas  
Q9f5t2 pseudomonas  
Q9f5t1 pseudomonas  
Q9f5t0 pseudomonas  
Q9f5s6 pseudomonas  
Q9f5t7 pseudomonas  
Q9f5s2 pseudomonas  
Q9zhm3 burkholderi  
Q8vum6 ralstonia s  
Q8vr23 pseudomonas  
Q8vl21 pseudomonas  
Q8vr24 pseudomonas  
Q8vr22 pseudomonas  
Q93na8 burkholderi  
Q8vr25 burkholderi  
Q8wrg8 alcaligenes  
Q9f5b6 marinobacte  
Q8vum5 naphthalene  
Q8vr21 burkholderi  
Q8vum2 ralstonia s  
Q93m40 uncultured  
Q93m39 uncultured  
Q93m41 uncultured

ALIGNMENTS

```
Db 121 KELYGESLNKKCLGLKEVARVESFHGFIYGCDFQAPSLMDYLGDAAWYLEPIFKHSGGL 180
QY 181 ELVGPPGKVIKANWKAPEAFNVGDYHVGWTHASSLSGESIFSSLAGNAALPPEGAGL 240
Db 181 ELVGPPGKVIKANWKAPEAFNVGDYHVGWTHASSLSRTGESIFSSLAGNAALPPEGAGL 240
QY 241 QMTSKYSGMGVLDWGYSGVHSADLVPELMAFGGAKQERLKEIGDVRARIYRSHLNCTV 300
Db 241 QMTSKYSGMGVLDWGYSGVHSADLVPELMAFGGAKQERLKEIGDVRARIYRSHLNCTV 300
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVTWYAIIVEKMDPEDLKRRLADSVQRTVGPAGFWES 360
Db 301 FPNNSVLTCSGVFKVWNPIDANTTEVTWYAIIVEKMDPEDLKRRLADAVQRTFGPAGFWES 360
QY 361 DDNDNMETASONGKKYQSRDSDLLSLNGLGFGEDVYGDVYPCGVGKSAIGETSYRGFYRAY 420
Db 361 DDNDNMETASONGKKYQSRDSDLLSLNGLGFGEDVYGDVYPCGVGKSAIGETSYRGFYRAY 420
QY 421 QAHVSSSSNWAEEFEDASSTWHTLTKTDDR 449
Db 421 QAHVSSSSNWAEEFEDASSTWHTLTKTDDR 449
```

## RESULT 2

```
Q52142
ID Q52142 PRELIMINARY; PRT; 449 AA.
AC Q52142: Q08194;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Iron-sulfur protein large subunit.
GN PAHAC.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA TAKIZAWA N.;
RA Kiyohara H.;
RT "Identification and characterization of genes encoding polycyclic aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon dihydrodiol dehydrogenase in Pseudomonas putida OUS82.";
RL J. Bacteriol. 176:2444-2449(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA Noboru T., Toshiya I., Takashi S., Kazuhiro Y., Masamichi K., Yue-Wu W., Masao F., Hohzoh K.;
RT "The molecular analysis of NAH7-type cluster located on the chromosomes of Pseudomonas aeruginosa Fsk1 and Pseudomonas putida OUS82.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA EMBL; AB004059; BAA20391.1; -.
DR HSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR Pfam; PF00355; Rieske; 1.
DR PRINTS; PR00048; Ring_hydroxyl_A; 1.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 449 AA; 49361 MW; 80020F54AAB1E8A CRC64;
```

Query Match 95.1%; Score 2291; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 4.7e-168;

```
Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;
QY 1 MMYNNKILVSESGLSQKHLIHGDEELFQHELKTTIFARNWLFLTHDSLIPAPGDYVTAKMG 60
Db 1 MMYNNKILVSESGUTQKHLIHGGGLFQHEURAVFARNWLFTHDSLIPSPGDYVTAKMG 60
QY 61 IDEVIVSRQNDGSIARAFILNVCRRHGKTLVSVYAGNAKGFVCSYHGWGFGSNGELQSPFPE 120
Db 61 IDEVIVSRQSDGSIARAFILNVCRRHGKTLVNAEAGNAKGFVCSYHGWGFGSNGELQSPFPE 120
QY 121 KDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQAPSLMDYLGDAAWYLEPMPFKHSGGL 180
Db 121 KELYGESLNKKCLGLKEVARVESFHGFIYACIDQEAFLMDYLGDAAWYLEPIFKHSGGL 180
QY 181 ELVGPPGKVIKANWKAPEAFNVGDYHVGWTHASSLSRSGESIFSSLAGNAALPPEGAGL 240
Db 181 ELVGPPGKVIKANWKAPEAFNVGDYHVGWTHASSLCTGESIFSSLAGNAALPPEGAGL 240
QY 241 QMTSKYSGMGVLDWGYSGVHSADLVPELMAFGGAKQERLKEIGDVRARIYRSHLNCTV 300
Db 241 QMTSKYSGMGVLDWGYSGVHSADLVPELMAFGGAKQERLKEIGDVRARIYRSHLNCTV 300
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVTWYAIIVEKMDPEDLKRRLADSVQRTVGPAGFWES 360
Db 301 FPNNSVLTCSGVFKVWNPIDANTTEVTWYAIIVEKMDPEDLKRRLADAVQRTFGPAGFWES 360
QY 361 DDNDNMETASONGKKYQSRDSDLLSLNGLGFGEDVYGDVYPCGVGKSAIGETSYRGFYRAY 420
Db 361 DDNDNMETASONGKKYQSRDSDLLSLNGLGFGEDVYGDVYPCGVGKSAIGETSYRGFYRAY 420
QY 421 QAHVSSSSNWAEEFEDASSTWHTLTKTDDR 449
Db 421 QAHVSSSSNWAEEFEDASSTWHTLTKTDDR 449

RESULT 3
Q92I73
ID Q92I73 PRELIMINARY; PRT; 449 AA.
AC Q92I73;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase Fe-S large subunit.
GN NAHAC.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AN10;
RX MEDLINE=99365311; PubMed=10433976;
RA Bosch R., Garcia-Valdes E., Moore E.R.B.;
RT "Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from Pseudomonas stutzeri AN10.";
RL Gene 236:149-157(1999).
DR EMBL; AF039533; AAD02136.1; -.
DR HSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR Pfam; PF00355; Rieske; 1.
DR PRINTS; PR00048; Ring_hydroxyl_A; 1.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 449 AA; 49802 MW; BA510FD1B6F56E63 CRC64;
```

Query Match 91.7%; Score 2207; DB 2; Length 449;  
Best Local Similarity 89.5%; Pred. No. 1.4e-161;  
Matches 402; Conservative 25; Mismatches 22; Indels 0; Gaps 0;

```
QY 1 MMYNNKILVSESGLSQKHLIHGDEELFQHELKTTIFARNWLFLTHDSLIPAPGDYVTAKMG 60
```



```
Db 1 MIYEN--LVSEAGLQKHLIHGDKELFQHELKTI FARNWLFLTHDSLIPSPGDIYVKAKMG 58
QY 61 IDEVIVSRQNDGSI RALNVCRRHRGKTLVSV EAGNAKGFVCSYHGWGFGSGNGELQSPVPE 120
Db 59 VDEVIVSRQNDGSI RALNVCRRHRGKTLVHA EAGNAKGFVCSYHGWGFGSGNGELQSPVPE 118
QY 121 KDLYGESLNKKCLGLKEVARVESFHGFIYGC FDOEAPPLMDYLGDAAWYLRPMPFKHSGGL 180
Db 119 KELYGDITKKKCLGLKEVPRIEFSHGFIYGC FDEAPFLVDYLGDAAWYLEPTFKHSGGL 178
QY 181 ELVGPPGKVVIVKANKKAPAEFNVDAYHVGW THASSLSRSGESIFSSLAGNAALPEEGAGL 240
Db 179 ELVGPPGKVVIVKANKKAPAEFNVDAYHVGW THASSLSRSGSIFTPLAGNAMLPEEGAGL 238
QY 241 QMTSKYSGMGVLDWGYSGVHSADLVPELMA FGGAKQERLNKEIGDVRARIYRSHLNCV 300
Db 239 QMTSKYSGMGVLDWGYSGVHSADLVPEMMA FGGAKQERLNKEIGDVRARIYRSHLNCV 298
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYA IVEKDMPEDLKRRLLADSVORTVGPAGFWES 360
Db 299 FPNNSILTCSGVFKVWNPIDENTTEVWTYA IVEKDMPEDLKRRLLADAVORTFGPAGFWES 358
QY 361 DDNDNMETASQNGKKYQSRDSDLLSNLGFGE DVGDAVYPGVWGKSAIGETSYRGFYRAY 420
Db 359 DDNDNMETESQNAKKYQSSNSDLIANLGF GKDVGDECYPCGWAKSAIGETSYRGFYRAY 418
QY 421 QAHVSSNNWAEFEHASTWHTLTKTDDR 449
Db 419 QAHISSSNNAEFENTSRNWHTELTKTTDR 447

RESULT 6
Q8VUD4
ID Q8VUD4 PRELIMINARY; PRT; 447 AA.
AC Q8VUD4;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE Dntac.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R34;
RX MEDLINE=20254695; PubMed=10795678;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Properties of the trihydroxytoluene oxygenase from Burkholderia cepacia R34: an extradiol dioxygenase from the 2,4-dinitrotoluene pathway."
RL Arch. Microbiol. 173:86-90(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=R34;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Origins of the 2,4-dinitrotoluene pathway."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF169302; AAL50021.1; -.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING-HYDROXYL_ALPHA; UNKNOWN1.
SQ SEQUENCE 447 AA; 49468 MW; 2A5BF8558320275C CRC64;

Query Match 86.2%; Score 2075; DB 2: Length 447;
Best Local Similarity 83.7%; Pred. No. 1.9e-151;
Matches 376; Conservative 33; Mismatches 33; Indels 2; Gaps 1;
QY 1 MNYNNKILVSESGLSQKHLIHGDEELFQHELKTI FARNWLFLTHDSLIPAPGDYVTKMG 60
```

```
Db 1 MSYQN--LVSEAGLQKHLIYDKELFQHELKTI FARNWLFLTHDSLIPSPGDIYVKAKMG 58
QY 61 IDEVIVSRQNDGSI RALNVCRRHRGKTLVSV EAGNAKGFVCSYHGWGFGSGNGELQSPVPE 120
Db 59 VDEVIVSRQNDGSI RALNVCRRHRGKTI VDAEAGNAKGFVCSYHGWGFGSGNGELQSPVPE 118
QY 121 KDLYGESLNKKCLGLKEVARVESFHGFIYGC FDOEAPPLMDYLGDAAWYLRPMPFKHSGGL 180
Db 119 KELYGDIAKKKCLGLKEVPRIEFSHGFIYGC FDEAPPLIDYLGDAAWYLEPTFKHSGGL 178
QY 181 ELVGPPGKVVIVKANKKAPAEFNVDAYHVGW THASSLSRSGESIFSSLAGNAALPEEGAGL 240
Db 179 ELVGPPGKVVIVKANKKAPAEFNVDAYHVGW THASSILRAGSIFAPLAGNAMEPEEGAGL 238
QY 241 QMTSKYSGMGVLDWGYSGVHSADLVPELMA FGGAKQERLNKEIGDVRARIYRSHLNCV 300
Db 239 QMTTKYSGIGVLWDAYSGIQSDMVPMPMAFGG AKQERLNKEIGDVRARIYRSHLNCV 298
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYA IVEKDMPEDLKRRLLADSVORTVGPAGFWES 360
Db 299 FPNNSFLTCSGVFKVFPIDENTTEVWTYA IVEKDMPEDLARRLADAVORSVGPAGYVES 358
QY 361 DDNDNMETASQNGKKYQSRDSDLLSNLGFGE DVGDAVYPGVWGKSAIGETSYRGFYRAY 420
Db 359 DDNDNMCTLSQNAKKYQSSNSDLIADLGF GKDVGDECYPCGWVGKSAISETSYRGFYRAY 418
QY 421 QAHVSSNNWAEFEHASTWHTLTKTDDR 449
Db 419 QAHISSSNNAEFENTSRNWHTELTKTTDR 447

RESULT 7
P95564
ID P95564 PRELIMINARY; PRT; 447 AA.
AC P95564;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DE TSPalpha 2Nt.
DE NTDAC.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JS42;
RX MEDLINE=97128768; PubMed=8973308;
RA Parales J.V., Kumar A., Parales R.E., Gibson D.T.;
RT "Cloning and sequencing of the genes encoding 2-nitrotoluene dioxygenase from Pseudomonas sp. JS42."
RL Gene 181:57-61(1996)
DR EMBL; U49504; AAB40383.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING-HYDROXYL_ALPHA; 1.
SQ SEQUENCE 447 AA; 49485 MW; ICB0E223E528E3BD CRC64;

Query Match 85.3%; Score 2055; DB 2: Length 447;
Best Local Similarity 83.7%; Pred. No. 6.7e-150;
Matches 376; Conservative 35; Mismatches 36; Indels 2; Gaps 1;
QY 1 MNYNNKILVSESGLSQKHLIHGDEELFQHELKTI FARNWLFLTHDSLIPAPGDYVTKMG 60
Db 1 MSYQN--LVSEAGLQKHLIHGDKELFQHELKTI FARNWLFLTHDSLIPSPGDIYVKAKMG 58
QY 61 IDEVIVSRQNDGSI RALNVCRRHRGKTLVSV EAGNAKGFVCSYHGWGFGSGNGELQSPVPE 120
Db 59 VDEVIVSRQNDGSI RALNVCRRHRGKTLVH TEAGNAKGFVCSYHGWGFGSGNGELQSPVPE 118
```



```
QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDAEAPPLMDYLGDAAWYLEPWFKHSGGL 180
Db 119 KELYGDAIKKCLGLKEVPRIESFHGFIYGCDFQDAEAPPLIDYLGDAAWYLEPWFKHSGGL 178
QY 181 ELVGPFGKVVIVKANWKAAPENFVGDAHYHGWTHASSLSRSGESFSSLAGNAALPPGAGL 240
Db 179 ELVGPFGKVVIVKANWKAAPENFVGDIYHGWTHAAALRAGQSVFSSLAGNAALPPGAGL 238
QY 241 QMTSKYSGMGVLWDGYSVHGSADLPVLMFAGGAQERLNKEIGDVRARIYRSHLNCV 300
Db 239 QMTSKYSGMGVLWDGYSVHGSADLPVLMFAGGAQERLNKEIGDVRARIYRSHLNCV 298
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMPEDLKRRLLADSVORTVGPAGFWES 360
Db 299 FPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMPEDLKRRLLADSVORTVGPAGFWES 358
QY 361 DDNDNMETASONGKKYQSRDLSNLGEGEDVYGDVAVPGVVGKSAIGTSGYRGFYRAY 420
Db 359 DDNDNMETASONGKKYQSRDLSNLGEGEDVYGDVAVPGVVGKSAIGTSGYRGFYRAY 418
QY 421 QAHVSSNNWAEFEHASSTWHTLTKTTDR 449
Db 419 QAHVSSNNWAEFEHASSTWHTLTKTTDR 447

RESULT 8
ID Q8RTL4 PRELIMINARY; PRT; 447 AA.
AC Q8RTL4;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Oxygenase-alpha NBDO.
GN NBZAC.
OS Comamonas sp. JS765.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=58226;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JS765;
RX MEDLINE=21681044; PubMed=11823201;
RA Lessner D.J., Johnson G.R., Perales R.E., Spain J.C., Gibson D.T.;
RT "Molecular Characterization and Substrate Specificity of Nitrobenzene
Dioxygenase from Comamonas sp. Strain JS765."
RL Appl. Environ. Microbiol. 68:634-641(2002).
DR EMBL; AF379638; AAL76202.1; .
SQ SEQUENCE 447 AA; 49556 MW; 4CA961F99A37DF3A CRC64;
```

```
Query Match 84.1%; Score 2026; DB 2; Length 447;
Best Local Similarity 82.0%; Pred. No. 1.1e-147;
Matches 368; Conservative 41; Mismatches 38; Indels 2; Gaps 1;

QY 1 MNYNNKILVSEGLSOKHLIHGDEELFQHELKTIIFARNWLFLTHDSLIPAPGDYVTAKWG 60
Db 1 MSYQN--LVSEAGLTOKHLIYGDKELFQHELKTIIFARNWLFLTHDSLIPSPGDYVKAKWG 58

QY 61 IDEVIVSRONDGSIRAFNVCRHRGKTLVSVAGNAKGFVCSYHGVGFGSNGELQSVPE 120
Db 59 VDEVIVSRONDGSVRAFLNVCRHRGKTLVHAEAGNAKGFVCGYHGVGFGSNGELQSVPE 118

QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDAEAPPLMDYLGDAAWYLEPWFKHSGGL 180
Db 119 KELYGDAIKKCLGLKEVPRIESFHGFIYGCDFQDAEAPPLIDYLGDAAWYLEPWFKHSGGL 178

QY 181 ELVGPFGKVVIVKANWKAAPENFVGDAHYHGWTHASSLSRSGESFSSLAGNAALPPGAGL 240
Db 179 ELVGPFGKVVIVKANWKAAPENFVGDIYHGWTHAAALRAGQSVFSSLAGNAALPPGAGL 238

QY 241 QMTSKYSGMGVLWDGYSVHGSADLPVLMFAGGAQERLNKEIGDVRARIYRSHLNCV 300
Db 239 QMTSKYSGMGVLWDGYSVHGSADLPVLMFAGGAQERLNKEIGDVRARIYRSHLNCV 298
```

```
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMPEDLKRRLLADSVORTVGPAGFWES 360
Db 299 FPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMPEDLKRRLLADSVORTVGPAGFWES 358
QY 361 DDNDNMETASONGKKYQSRDLSNLGEGEDVYGDVAVPGVVGKSAIGTSGYRGFYRAY 420
Db 359 DDNDNMETASONGKKYQSRDLSNLGEGEDVYGDVAVPGVVGKSAIGTSGYRGFYRAY 418
QY 421 QAHVSSNNWAEFEHASSTWHTLTKTTDR 449
Db 419 QAHVSSNNWAEFEHASSTWHTLTKTTDR 447

RESULT 9
ID Q45695 PRELIMINARY; PRT; 451 AA.
AC Q45695;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ISP-alpha.
GN DNTAC.
OS Burkholderia sp. (strain RASC).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=69003;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DNT;
RX MEDLINE=96326338; PubMed=8759857;
RA Suen W.C., Haigler B.E., Spain J.C.;
RT "2,4-Dinitrotoluene dioxygenase from Burkholderia sp. strain DNT:
similarity to naphthalene dioxygenase."
RT J. Bacteriol. 178:4926-4934(1996).
DR EMBL; U62430; AAB09766.1; .
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 451 AA; 49827 MW; 726796C29CEF9A10 CRC64;

Query Match 81.7%; Score 1968.5; DB 2; Length 451;
Best Local Similarity 80.0%; Pred. No. 3.1e-143;
Matches 359; Conservative 41; Mismatches 46; Indels 3; Gaps 2;

QY 1 MNYNNKILVSEGLSOKHLIHGDEELFQHELKTIIFARNWLFLTHDSLIPAPGDYVTAKWG 60
Db 6 MSYQN--LVSEAGLTOKHLIYGDKELFQHELKTIIFARNWLFLTHDSLIPSPGDYVKAKWG 63

QY 61 IDEVIVSRONDGSIRAFNVCRHRGKTLVSVAGNAKGFVCSYHGVGFGSNGELQSVPE 120
Db 64 VDEVIVSRONDGSVRAFLNVCRHRGKTIIVDAEAGNAKGFVCGYHGVGFGSNGELQSVPE 123

QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDAEAPPLMDYLGDAAWYLEPWFKHSGGL 180
Db 124 KELYGDAIKKCLGLKEVPRIESFHGFIYGCDFQDAEAPPLIDYLGDAWYLEPWFKHSGGL 183

QY 181 ELVGPFGKVVIVKANWKAAPENFVGDAHYHGWTHASSLSRSGESFSSLAGNAALPPGAGL 240
Db 184 ELVGPFGKVVIVKANWKAAPENFVGDIYHGWTHASSLSRSGESFSSLAGNAALPPGAGL 243

QY 241 QMTSKYSGMGVLWDGYSVHGSADLPVLMFAGGAQERLNKEIGDVRARIYRSHLNCV 300
Db 244 QMTSKYSGMGVLWDGYSVHGSADLPVLMFAGGAQERLNKEIGDVRARIYRSHLNCV 303

QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMPEDLKRRLLADSVORTVGPAGFWES 360
Db 304 FPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMPEDLKRRLLADSVORTVGPAGFWES 363

QY 361 DDNDNMETASONGKKYQSRDLSNLGEGEDVYGDVAVPGVVGKSAIGTSGYRGFYRAY 420
```

```
||||| ||| |||| :|||:||||:|||| |||| ||| |||:|||||
Db 364 DDNDNK-VLSNAKKYQSSDLADLGKDYGDCEYGVYKSAFSETNHRGFYAY 422
QY 421 QAHVSSNAEFEEHASSTWHTLTKTTDR 449
|||:|||||: | |||||
Db 423 QAHISSSNAEFENTSRNWHTELTKTTDR 451

RESULT 10
Q9ETK2 PRELIMINARY; PRT; 277 AA.
AC Q9ETK2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2 OR NAHAC.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2IDINH, PR1M1, AND 31A2NH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306438; AAG25698.1; -
DR EMBL; AF306432; AAG25692.1; -
DR EMBL; AF306436; AAG25696.1; -
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PRO0090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 30094 MW; BFCF9B5A379546DB CRC64;

Query Match 61.9%; Score 1490; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1e-106;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRQDGSIRAFNLVCRHRGKTLVSVBAGNAKGFVCSYHGWGFGSNGELQ 115
|||||
Db 1 TAKMGIDEIVSRQDGSIRAFNLVCRHRGKTLVSVBAGNAKGFVCSYHGWGFGSNGELQ 60
QY 116 SVPFKDLYGESLNKCLGLKEVARVESFHGFIYGCDFQDQAPPLMDYLGDAAWYLEPMFK 175
|||||
Db 61 SVPFKDLYGESLNKCLGLKEVARVESFHGFIYGCDFQDQAPPLMDYLGDAAWYLEPMFK 120
QY 176 HSGGLELVGPPGVKVIKANKKAPAEFVGDYHVGWTHASSLSRGSIFSSLAGNAALPP 235
|||||
Db 121 HSGGLELVGPPGVKVIKANKKAPAEFVGDYHVGWTHASSLSRGSIFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPELMAFGGAKQERLNKEIGDVRARIYRSH 295
|||||
Db 181 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPELMAFGGAKQERLNKEIGDVRARIYRSH 240
QY 296 LNCVTFPNNMLTCSGVFKVWNPIDANTTEVWTYAIY 332
|||||
Db 241 LNCVTFPNNMLTCSGVFKVWNPIDANTTEVWTYAIY 277

RESULT 11
Q9F5S4 PRELIMINARY; PRT; 277 AA.
ID Q9F5S4
AC Q9F5S4;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
```

```
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5IIIASAL;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306440; AAG25700.1; -
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PRO0090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 30094 MW; FADFAD5D37E241AB CRC64;

Query Match 61.8%; Score 1488; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 1.4e-106;
Matches 276; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRQDGSIRAFNLVCRHRGKTLVSVBAGNAKGFVCSYHGWGFGSNGELQ 115
|||||
Db 1 TAKMGIDEIVSRQDGSIRAFNLVCRHRGKTLVSVBAGNAKGFVCSYHGWGFGSNGELQ 60
QY 116 SVPFKDLYGESLNKCLGLKEVARVESFHGFIYGCDFQDQAPPLMDYLGDAAWYLEPMFK 175
|||||
Db 61 SVPFKDLYGESLNKCLGLKEVARVESFHGFIYGCDFQDQAPPLMDYLGDAAWYLEPMFK 120
QY 176 HSGGLELVGPPGVKVIKANKKAPAEFVGDYHVGWTHASSLSRGSIFSSLAGNAALPP 235
|||||
Db 121 HSGGLELVGPPGVKVIKANKKAPAEFVGDYHVGWTHASSLSRGSIFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPELMAFGGAKQERLNKEIGDVRARIYRSH 295
|||||
Db 181 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPELMAFGGAKQERLNKEIGDVRARIYRSH 240
QY 296 LNCVTFPNNMLTCSGVFKVWNPIDANTTEVWTYAIY 332
|||||
Db 241 LNCVTFPNNMLTCSGVFKVWNPIDANTTEVWTYAIY 277

RESULT 12
Q9F5T3 PRELIMINARY; PRT; 277 AA.
ID Q9F5T3
AC Q9F5T3;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 19IIDNH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139738;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=19IIDNH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AF306426; AAG25686.1; -.
DR HSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30064 MW; DB47868EB6DD525A CRC64;

Query Match 61.7%; Score 1486; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 2.1e-106;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRNDGSIKRAFLNVCVRHGRKTLVSVVAGNAKGFVCSYHGWGFGSNGELQ 115
DB 1 TAKMGIDEVIVSRNDGSIKRAFLNVCVRHGRKTLVSVVAGNAKGFVCSYHGWGFGSNGELQ 60

QY 116 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQEQAPPLMDYLGDAAWYLEPMFK 175
DB 61 GVPPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQEQAPPLMDYLGDAAWYLEPMFK 120

QY 176 HSGGLELVGPPGKVVIRKANWKAPAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVIRKANWKAPAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPP 180

QY 236 EGAGLQMTSKYSGMGVLDGYSVHSDLVPELMFAGGAKQERLNKEIGDVRARIYRSH 295
DB 181 EGAGLQMTSKYSGMGVLDGYSVHSDLVPELMFAGGAKQERLNKEIGDVRARIYRSH 240

QY 296 LNCTVFPNNSMLTCGVSFKVWNPIDANTTEWTYAI 332
DB 241 LNCTVFPNNSMLTCGVSFKVWNPIDANTTEWTYAI 277

RESULT 13
Q9F5S8 PRELIMINARY; PRT; 277 AA.
AC Q9F5S8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Naphthalene dioxxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. PR3MN2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR3MN2;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306434; AAG25694.1; -.
DR HSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30064 MW; 6BFF83D32F04CFC3 CRC64;

Query Match 61.6%; Score 1484; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 2.9e-106;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRNDGSIKRAFLNVCVRHGRKTLVSVVAGNAKGFVCSYHGWGFGSNGELQ 115
DB 1 TAKMGIDEVIVSRNDGSIKRAFLNVCVRHGRKTLVSVVAGNAKGFVCSYHGWGFGSNGELQ 60

QY 116 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQEQAPPLMDYLGDAAWYLEPMFK 175
DB 61 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQEQAPPLMDYLGDAAWYLEPMFK 120

QY 176 HSGGLELVGPPGKVVIRKANWKAPAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVIRKANWKAPAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPP 180

QY 236 EGAGLQMTSKYSGMGVLDGYSVHSDLVPELMFAGGAKQERLNKEIGDVRARIYRSH 295
DB 181 EGAGLQMTSKYSGMGVLDGYSVHSDLVPELMFAGGAKQERLNKEIGDVRARIYRSH 240

QY 296 LNCTVFPNNSMLTCGVSFKVWNPIDANTTEWTYAI 332
DB 241 LNCTVFPNNSMLTCGVSFKVWNPIDANTTEWTYAI 277

RESULT 14
Q9F5S9 PRELIMINARY; PRT; 277 AA.
AC Q9F5S9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Naphthalene dioxxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 8IDINH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8IDINH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306433; AAG25693.1; -.
DR HSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 29995 MW; AADA9B5A378FF6DB CRC64;

Query Match 61.6%; Score 1483; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 3.5e-106;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRNDGSIKRAFLNVCVRHGRKTLVSVVAGNAKGFVCSYHGWGFGSNGELQ 115
DB 1 TAKMGIDEVIVSRNDGSIKRAFLNVCVRHGRKTLVSVVAGNAKGFVCSYHGWGFGSNGELQ 60

QY 116 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQEQAPPLMDYLGDAAWYLEPMFK 175
DB 61 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQEQAPPLMDYLGDAAWYLEPMFK 120

QY 176 HSGGLELVGPPGKVVIRKANWKAPAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVIRKANWKAPAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPP 180

QY 236 EGAGLQMTSKYSGMGVLDGYSVHSDLVPELMFAGGAKQERLNKEIGDVRARIYRSH 295
```

```
|||||
Db 181 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPDLMAFGAKQEGLNKEIGDVRARIYRSH 240
QY 296 LNCCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 332
Db 241 LNCCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 277

RESULT 15
Q9F5S3 PRELIMINARY; PRT; 277 AA.
AC Q9F5S3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 1 (fragment).
GN NAHAc1.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=51IANH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahAc) in
RT Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308441; AAG25701.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PRO0090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 3062 MW; 82C78D865A92D0AB CRC64;

Query Match 60.3%; Score 1451; DB 2; Length 277;
Best Local Similarity 96.8%; Pred. No. 1e-103;
Matches 268; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRQDGSIRAFNLNCRHRGKTLVSVEAGNAKGFVCSYHGWFSGSNGELQ 115
Db 1 TAKMGIDEVIVSRQDGSIRAFNLNCRHRGKTLVNAEAGNAKGFVCSYHGWFSGSNGELQ 60

QY 116 SVPFKDLYGESLNKKLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFK 175
Db 61 SVPFKELYGESLNKKLGLKEVARVESFHGFIYGCDFQDEAPSLMDYLGDAAWYLEPIFK 120

QY 176 HSGGLELVGPPGKVKVVIKANWKAPAEFNGDAYHVGWTHASSLRSGESIFSSLAGNAALPP 235
Db 121 HSGGLELVGPPGKVKVVIKANWKAPAEFNGDAYHVGWTHASSLRSGESIFSSLAGNAVLPP 180

QY 236 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPDLMAFGAKQERLNKEIGDVRARIYRSH 295
Db 181 EGAGLQMTSKYSGMGVLWDGYSGVHSADLVPDLMAFGAKQERLNKEIGDVRARIYRSH 240

QY 296 LNCCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 332
Db 241 LNCCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 277
```

Search completed: January 27, 2003, 09:02:17  
Job time : 27.8333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:47:32 ; Search time 7.16667 Seconds  
(without alignments)  
2598.540 Million cell updates/sec

Title: US-09-843-250-2  
Perfect score: 2408  
Sequence: 1 MNYNNKILVSEGLSOKHLI.....AFFEHASSTWHTLTAKTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2403	99.8	449	1	NDOB_PSEPU	P23094 pseudomonas
2	2317	96.2	449	1	NDOB_PSEFL	O07824 pseudomonas
3	2218	92.1	449	1	NDOB_PSEAE	O51494 pseudomonas
4	736	30.6	453	1	HCAE_ECOLI	O47139 escherichia
5	651.5	27.1	450	1	BEDI_PSEPU	O07944 pseudomonas
6	649.5	27.0	450	1	TODI_PSEPU	P13450 pseudomonas
7	616.5	25.6	458	1	BPHA_BURCE	P37333 burkholderi
8	616	25.6	457	1	BPHA_COMTE	O46372 comamonas t
9	616	25.6	458	1	BPHA_PSEPS	O52028 pseudomonas
10	588.5	24.4	448	1	BNZA_PSEPU	P08084 pseudomonas
11	586	24.3	458	1	BPAL_PSESI	O52438 pseudomonas
12	410.5	17.0	454	1	XYLK_PSEPU	P23099 pseudomonas
13	398	16.5	461	1	BENA_ACICA	P07769 acinetobact
14	248	10.3	374	1	YEAN_ECOLI	P76253 escherichia
15	239.5	9.9	442	1	CHMO_AMATR	O93xel1 amaranthus
16	233	9.7	438	1	CHMO_ATRHO	O91kn0 atriplex ho
17	224	9.3	436	1	CHMO_BETVU	O22553 beta vulgar
18	223	9.3	439	1	CHMO_SPIOL	O04121 spinacia ol
19	216	9.0	422	1	CHMO_ARATH	O9szr0 arabidopsis
20	163	6.8	439	1	PHY3_PSEPU	O05183 pseudomonas
21	118.5	4.9	354	1	VANA_PSESP	O05616 pseudomonas
22	112	4.7	1276	1	PMF6_CHLPN	O9z899 chlamydia p
23	110	4.6	329	1	VANA_PSES9	P12609 pseudomonas
24	110	4.6	432	1	CBAA_COMTE	O44256 comamonas t
25	103	4.3	409	1	POBA_PSEPS	O52185 pseudomonas
26	101.5	4.2	3255	1	POLG_LMVE	P89876 l genome po
27	100.5	4.2	543	1	TUPL_DROME	P16375 drosophila
28	100.5	4.2	746	1	TUPL_DROME	P16376 drosophila
29	100.5	4.2	1926	1	LPH_RABIT	P09849 oryctolagus
30	99.5	4.1	468	1	PPAL_PICPA	P52291 pichia past
31	98.5	4.1	1411	1	Y297_HUMAN	O15040 homo sapien
32	97.5	4.0	3255	1	POLG_LMVO	P31999 l genome po
33	94.5	3.9	331	1	LDHA_RHIDE	O9pw58 rhigophila

RESULT 1									
NDOB_PSEPU									
ID	NDOB_PSEPU	STANDARD;	350	1	UNRI_HUMAN	Q9y3f4	homo sapien		
AC	P23094; O52124; O33461; O07830;	PRT; 449 AA.	350	1	F26_YEAST	P32604	saccharomyc		
DT	01-NOV-1991 (Rel. 20, Created)		452	1	ATPA_ENTHR	P26679	enterococcu		
DT	01-NOV-1991 (Rel. 20, Last sequence update)		518	1	ORP8_HUMAN	Q9bzf1	homo sapien		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		847	1	SYL_STAAU	P41972	staphylococ		
DE	Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene 1,2-dioxygenase ISP alpha).		917	1	DHE3_PYRKO	O59650	pyrococcus		
DE	NDOB OR NAHAC OR DOXB OR NAHA3 OR NDOC2.		420	1	POLN_HEVMY	O04610	hepatitis e		
OS	Pseudomonas putida, and		1693	1	GBAL_CANAL	P28868	candida alb		
OG	Flasmid pDGI, Plasmid NAH7, and Plasmid NPL1.		849	1	DLG3_MOUSE	P70175	mus musculu		
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;		849	1	DCP2_PEA	P51851	pisum sativ		
OC	Pseudomonas.		405	1	AGLU_ASFOR	Q12558	aspergillus		
OX	NCBI_TaxID=303, 306;		3.8						
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	SPECIES=P.putida; STRAIN=NCIB 9816;								
RX	MEDLINE=89211973; PubMed=3243438;								
RA	Kurkela S., Leivaeslaih H., Palva E.T., Teeri T.H.;								
RT	"Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of Pseudomonas putida strain NCIB9816.";								
RL	Gene 73:355-362(1988).								
RP	SEQUENCE FROM N.A.								
RC	SPECIES=P.putida; STRAIN=NCIB 9816;								
RX	Parales J.V., Parales R.E., Kumar A., Gibson D.T.;								
RA	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=C18;								
RX	MEDLINE=94042852; PubMed=8226631;								
RA	Denome S.A., Stanley D.C., Olson E.S., Young K.D.;								
RT	"Metabolism of dibenzothiophene and naphthalene in Pseudomonas strains: complete DNA sequence of an upper naphthalene catabolic pathway.";								
RL	J. Bacteriol. 175:6890-6901(1993).								
RP	SEQUENCE FROM N.A.								
RC	SPECIES=P.putida; STRAIN=G7 / ATCC 17485; PLASMID=NAH7;								
RX	MEDLINE=93552277; PubMed=8486285;								
RA	Simon M.J., Oslund T.D., Saunders R., Ensley B.D., Suggs S., Harcourt A.A., Suen W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;								
RT	"Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida strains G7 and NCIB 9816-4.";								
RL	Gene 127:31-37(1993).								
RP	SEQUENCE FROM N.A.								
RC	SPECIES=P.putida; STRAIN=BS202; PLASMID=NPL1;								
RA	Bezborodnikov S.G., Boronin A.M., Tiedje J.M.;								
RT	"Nucleotide sequences of genes encoding an upper pathway of naphthalene metabolism of NPL1 plasmid from Pseudomonas putida strain BS202.";								
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.								



```
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF004283; AB61370.1; -.
CC HSSP; P23094; INDO.
CC InterPro; IPR001281; Rieske.
CC InterPro; IPR001663; Ring_hydroxyl_A.
CC Pfam; PF00335; Rieske; 1.
CC Pfam; PF00848; Ring_hydroxyl_A; 1.
CC PRINTS; PR00090; RINGDIOXGNASE.
CC PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC DIOXYGENASE; NAD; Plasmid.
CC METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 208 208 IRON (BY SIMILARITY).
CC METAL 213 213 IRON (BY SIMILARITY).
CC METAL 362 362 IRON (BY SIMILARITY).
CC SEQUENCE 449 AA; 49512 MW; 186A7924AD437A1C CRC64;
Query Match 96.2%; Score 2317; DB 1; Length 449;
Best Local Similarity 95.5%; Pred. No. 1.5e-174;
Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;
QY 1 MNYNKKLVSESGLSQKHLHGDEELFQHELKTFIFARNWFLTHDSLIPAGDYVTAKMG 60
DB 1 MNYNKKLVSESGLSQKHLHGDEELFQHELKTFIFARNWFLTHDSLIPAGDYVTAKMG 60
QY 61 IDEVIVSRQDGSIRAFNLVNCVRHGKTLVSVAGNAKGFVCSYHGMFGSGNGELQSVPE 120
DB 61 IDEVIVSRQDGSIRAFNLVNCVRHGKTLVNAEAGNAKGFVCSYHGMFGSGNGELQSVPE 120
QY 121 KDLYGESLNKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLDGAAWYLEPFIKHSGL 180
DB 121 KELYGESLNKKCLGLKEVARVESFHGFIYGCFOEAPSLMDYLDGAAWYLEPFIKHSGL 180
QY 181 ELYGPGKVVIKANWKAPAEFVGDAYHVGWTHASSLSGSEIFSSLAGNAALPPGAGL 240
DB 181 ELYGPGKVVIKANWKAPAEFVGDAYHVGWTHASSLSGSEIFSSLAGNAALPPGAGL 240
QY 241 QMTSKYSGMGVLWDGYSGVHSDLPVLPFMAFGAKQERLNKEIGDVRARIYRSHLNCTV 300
DB 241 QMTSKYSGMGVLWDGYSGVHSDLPVLPFMAFGAKQERLNKEIGDVRARIYRSHLNCTV 300
QY 301 FPNNSMLTCSGVFKWNPIDANTTEVWYTAIVEKDPEDLKRRLADSVQRTGVPAGFWS 360
DB 301 FPNNSMLTCSGVFKWNPIDANTTEVWYTAIVEKDPEDLKRRLADSVQRTGVPAGFWS 360
QY 361 DDNDNMTASQNGKKYQSRDSDLLSNLFGEDYGDVAVPGVVGKSAIGTSTRGFYRAY 420
DB 361 DDNDNMTASQNGKKYQSRDSDLLSNLFGEDYGDVAVPGVVGKSAIGTSTRGFYRAY 420
Db 361 DDNDNMTASQNGKKYQSRDSDLLSNLFGEDYGDVAVPGVVGKSAIGTSTRGFYRAY 420
QY 421 QAHVSSSNWAEFEHASSSTWHTLTKTTR 449
Db 421 QAHVSSSNWAEFEHASSSTWHTLTKTTR 449
Db 421 QAHVSSSNWAEFEHASSSTWHTLTKTTR 449
RESULT 3
NDOB_PSEAE
ID NDOB_PSEAE STANDARD; PRT; 449 AA.
AC Q51494;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene
DE 1,2-dioxygenase ISP alpha).
GN NDOB OR PAHA3.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PaK1;
RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,
RA Kiyohara H.;
RT "The molecular analysis of an NAH7-type gene cluster, pah, located on
RT the chromosome of Pseudomonas aeruginosa PaK1."
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO).
CC BOTH COMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL.
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D84146; BAA12240.1; -.
CC HSSP; P23094; INDO.
CC InterPro; IPR001281; Rieske.
CC InterPro; IPR001663; Ring_hydroxyl_A.
CC Pfam; PF00335; Rieske; 1.
CC Pfam; PF00848; Ring_hydroxyl_A; 1.
CC PRINTS; PR00090; RINGDIOXGNASE.
CC PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC DIOXYGENASE; NAD.
CC METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 208 208 IRON (BY SIMILARITY).
```

```
FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 362 362 IRON (BY SIMILARITY).
SQ SEQUENCE 449 AA; 49715 MW; 35A18913672A21C CRC64;

Query Match 92.1%; Score 2218; DB 1; Length 449;
Best Local Similarity 89.8%; Pred. No. 8.8e-167;
Matches 403; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

QY 1 MMYNKILVSEGLSQHLIHGDELFQHELKTIFARNWLFTHDSLIPAGDYVTAKMG 60
Db 1 MMYKNLVSEGLQKHLIHGDELFQHELKTIFARNWLFTHDSLIPAGDYVTAKMG 60
QY 61 IDEVIVSRQNDGSIKRAFLNVCVRHCKTILVSVFAGNAKGFVCSYHGWGFGANGELQSPFPE 120
Db 61 VDEVIVSRQNDGSIKRAFLNVCVRHCKTILVSVFAGNAKGFVCSYHGWGFGANGELQSPFPE 120
QY 121 KDLYESLNKKGCLKEVARVESFHGTYGCFDQAPPLMDYLDGAAWYLPMPKHSGL 180
Db 121 KELYEALDKKCMGLKEVARVESFHGTYGCFDQAPPLMDYLDGAAWYLPMPKHSGL 180
QY 181 ELVGGPGKVIKANKKAENFVGDYHVGWTHASSLSRSGESIFSSLAGNAALPEAGL 240
Db 181 ELVGGPGKVIKANKKAENFVGDYHVGWTHASSLSRSGESIFSSLAGNAALPEAGL 240
QY 241 QMTSYGSGMGLDGYSGVHSADLVPELMAFGGAKQERLKEICDVRARIYRSHLNCTV 300
Db 241 QMTSYGSGMGLDGYSGVHSADLVPELMAFGGAKQERLKEICDVRARIYRSHLNCTV 300
QY 301 FPNNSMLTCSGVKFWNPIDANTTEWTYAIIVEKDPEDLKRRLADSVQRTVGPAGFWS 360
Db 301 FPNNSFLTCSGVKFWNPIDANTTEWTYAMVEKDPEDLKRRLADSVQRTVGPAGFWS 360
QY 361 DNDNNETASQKKYQRSDLLNSLNLGFGEDYVDYGVGVGKSAIGETSYRGFYRAY 420
Db 361 DNDNNETVQNAKQYQRSDGLVNSLNLGFGEDYVDYGVGVGKSAIGETSYRGFYRAY 420
QY 421 QARVSSSSWAEPEHASTWHTLTKTDR 449
Db 421 GAHSSSSWAEFEDYSKNWHTLTKTDR 449

RESULT 4
HCAE_ECOLI STANDARD; PRT; 453 AA.
AC Q47139; P77590; P78203;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 3-phenylpropionate dioxygenase alpha subunit (EC 1.14.1.-)
DE (Digoxigenin alpha subunit).
GN HCAE OR PHD1 OR HCAA OR HCAAL OR DIGA OR B2538 OR Z3809 OR ECS3404.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-K12;
RA Turlin E., Gasser F., Biville F.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Oshibuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Ozima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
[4]
RN SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch K.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[5]
RN SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[6]
RN FUNCTION: PART OF MULTICOMPONENT 3-PHENYLPROPIONATE DIOXYGENASE.
RN CONVERTS 3-PHENYLPROPIONIC ACID (PP) INTO CIS-3-(3-CARBOXYETHYL)-
RN 3,5-CYCLOHEXADIENE-1,2-DIOL (PP-DIHYDRODIOL).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: 3-phenylpropionic acid catabolism.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (HCAE AND HCAF), A
CC FERREDOXIN (HCAC) AND A FERREDOXIN REDUCTASE (HCAD).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z37966; CAAB6018.1; -
DR EMBL; AE000340; AAC75591.1; -
DR EMBL; D90883; BAA16433.1; -
DR EMBL; D90884; BAA16441.1; -
DR EMBL; AE005484; AAG57651.1; -
DR EMBL; AF002562; BAB36827.1; -
DR HSSP; P23094; LINDO.
DR EcoGene; EGI3456; hcaE.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYLALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; 85 85
FT METAL 87 87 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 105 105 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
```





DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Toluene 2,3-dioxygenase alpha subunit (EC 1.14.12.-).  
 GN TODC1.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 ON NCBI\_TaxID=303;  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.  
 RC STRAIN=F1;  
 RX MEDLINE=89359301; PubMed=2670929;  
 RA Zylstra G.J., Gibson D.T.;  
 RT "Toluene degradation by Pseudomonas putida F1. Nucleotide sequence of  
 the todC1C2BADE genes and their expression in Escherichia coli.";  
 RL J. Biol. Chem. 264:14940-14946(1989).  
 CC -!- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
 CC -!- PATHWAY: Toluene degradation; first step.  
 CC -!- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE  
 TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (TODC1 AND TODC2), A  
 FERREDOXIN (TOD3) AND A FERREDOXIN REDUCTASE (TODA).  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 DIOXYGENASE ALPHA SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J04996; AAA26005.1; -  
 DR PIR; A36516; A36516.  
 DR HSSP; P23094; INDO.  
 DR InterPro; IPR001281; Rieske.  
 DR InterPro; IPR001663; Ring\_hydroxyl\_A.  
 DR Pfam; PF00355; Rieske; 1.  
 DR Pfam; PF00848; Ring\_hydroxyl\_A; 1.  
 DR PRINTS; PR00090; RINGDIOXGNASE.  
 DR PROSITE; PS00570; RING-HYDROXYL\_ALPHA; 1.  
 DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
 KW Dioxygenase; NAD.  
 FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 222 222 IRON (BY SIMILARITY).  
 FT METAL 228 228 IRON (BY SIMILARITY).  
 SQ SEQUENCE 450 AA; 50944 MW; 038C80F197F3485D CRC64;  
 Query Match 27.0%; Score 649.5; DB 1; Length 450;  
 Best Local Similarity 34.4%; Pred. No. 1.4e-43;  
 Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;  
 QY 20 IGDDELQHEHLTKTIFARNWFLFHDLSLPAGDYVYAKGIDVIVSRONGSIRAFNL 79  
 DB 35 IITDELQLELRFARSWLLLGHEQTIRKPGDYITYMGEDPVVVRORDASIAVFLN 94  
 QY 80 VCRHGRKTLVSVEAGNAGKFCVSYHGWMGFGSGNGELQSVPFKDLGSLNKKCGLGLKE-- 137  
 DB 95 QCRHGRMTRCADAGNAKAFCTSYHGWAYDTAGNLVNVPEAESFA-----CLNKKWS 148  
 QY 138 --VARVESEHFIYCCFQDQAPPMLDYLGDAAWYLEPMF-KHSGGLVGPQKVVTKAN 194  
 DB 149 PLKARVETKYGLIFANWDENAVDLDITLGEAKFYMDHMLDRTEAGTEAIPGVQKWVPCN 208  
 QY 195 WKAPAEVGDAYHVGWT-HASSLRSG--ESIFSSLAGNALPPEGAGLQMTSKYSGSMG 251  
 DB 209 WKFAAEQFCSDMYRAGTITSHLSILAGLPEDL-----EMADLAPPTVGKQYRASWGSHGS 263  
 QY 252 VLWDGYSGVHSADLVPELMAF--GGAKQERLNKEIGDVR--ARIYRSHLNCTVFPNNSML 307

DB 264 GFYVGDPLNLMLAIMGPKVTSYWTGPASERAAERLGSVERGSKLMVEHM--TVFPTCSFL 321  
 QY 308 TCSGVFKVWNPIDANTTEVWTYIAIVEKMDPEDLKRRLADSVQRTVGPAGFWESDDNDNME 367  
 DB 322 PGINTVTHPRGPNFNEVWAFVVDADAPDDIKEEFRQTLRTFSAGGVFEQDDGENWY 381  
 QY 368 TASONGKQYQSRDLSNLGFGEDVDVAVYPGVWGKSAIGETSYRGFYRAYQAHVSSS 427  
 DB 382 EIQHILRGHKARSPPFNAEMSDQTVNDPVPYFORISNNVYSEEAARGLYAHLMRMSTP 441  
 QY 428 NW 429  
 DB 442 DW 443  
 RESULT 7  
 BPHA\_BURCE  
 ID BPHA\_BURCE STANDARD; PRT; 458 AA.  
 AC P37333;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-  
 dioxygenase).  
 GN BPHA.  
 OS Burkholderia cepacia (Pseudomonas cepacia).  
 OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;  
 OC Burkholderia.  
 ON NCBI\_TaxID=292;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LB400;  
 RX MEDLINE=92234948; PubMed=1569021;  
 RA Erickson B.D., Mondello F.J.;  
 RT "Nucleotide sequencing and transcriptional mapping of the genes  
 encoding biphenyl dioxygenase, a multicomponent  
 polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain  
 LB400."  
 RT LB400."  
 RL J. Bacteriol. 174:2903-2912(1992).  
 RN [2]  
 RP SEQUENCE OF 1-10, AND CHARACTERIZATION.  
 RC STRAIN=LB400;  
 RX MEDLINE=96011369; PubMed=7592331;  
 RA Haddock J.D., Gibson D.T.;  
 RT "Purification and characterization of the oxygenase component of  
 biphenyl 2,3-dioxygenase from Pseudomonas sp. strain LB400."  
 RL J. Bacteriol. 177:5834-5839(1995).  
 RN [3]  
 RP ERRATUM.  
 RA Haddock J.D., Gibson D.T.;  
 RL J. Bacteriol. 178:2158-2158(1996).  
 CC -!- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-  
 phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).  
 CC -!- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit.  
 CC -!- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;  
 CC first step.  
 CC -!- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three  
 Bphe subunits. A ferredoxin (BphF) and a ferredoxin reductase  
 (BphG) must be present to obtain activity.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 DIOXYGENASE ALPHA SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M86348; AAB63425.1;  
 DR PIR; B41858; B41858.







[illegible]







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:55:37 ; Search time 13.1667 Seconds  
(without alignments)  
3278.305 Million cell updates/sec

Title: US-09-843-250-2

Perfect score: 2408

Sequence: 1 MNVNNKIIVSESGLSQKHLI.....AEFEHASTWHTLTKRTDR 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2403	99.8	449	2 JS0071	naphthalene dioxyg
2	2403	99.8	462	2 S27632	naphthalene 1,2-di
3	2345	97.4	449	2 JN0644	naphthalene 1,2-di
4	2291	95.1	449	2 C55217	polycyclic aromati
5	2055	85.3	447	2 JCS352	2-nitrotoluene dio
6	956	39.7	459	2 T31134	naphthalene dioxyg
7	773	32.1	450	2 T31256	terminal oxygenase
8	736	30.6	453	2 A65031	biphenyl dioxygena
9	736	30.6	453	2 D91054	biphenyl dioxygena
10	736	30.6	453	2 G85898	biphenyl dioxygena
11	724.5	30.1	455	2 T31258	aromatic oxygenase
12	651.5	27.1	450	1 JN0812	benzene 1,2-dioxyg
13	649.5	27.0	450	1 A36516	toluene dioxygens
14	638	26.5	461	2 S51757	biphenyl dioxygena
15	616.5	25.6	459	1 B41858	biphenyl dioxygena
16	616	25.6	457	1 JC4993	biphenyl dioxygena
17	616	25.6	458	1 A42409	biphenyl dioxygena
18	604	25.1	431	2 JMO098	carbazole dioxygen
19	588.5	24.4	448	1 A29830	benzene 1,2-dioxyg
20	586	24.3	458	2 JC2467	biphenyl dioxygena
21	451	18.7	469	2 T50934	dioxygenase DfAl,
22	417	17.3	471	2 T31281	benzoate 1,2-dioxy
23	410.5	17.0	454	2 A41659	toluate 1,2-dioxy
24	408.5	17.0	455	2 E83332	probable benzoate
25	393	16.3	461	2 S23477	antirrhizole dioxy
26	390	16.2	464	2 G83331	probable ring-hydr
27	340.5	14.1	424	2 E83384	biphenyl dioxygena
28	334	13.9	426	2 T31278	aromatic oxygenase
29	314.5	13.1	391	2 T31251	

30	311.5	12.9	468	2 G97447	hypothetical prote
31	311.5	12.9	468	2 AH2665	ring hydroxylating
32	295	12.3	429	2 F82970	probable ring hydr
33	287	11.9	412	2 H95311	probable aromatic
34	283	11.8	415	2 AG3320	benzoate 1,2-dioxy
35	270.5	11.2	404	2 H87635	Rieske 2Fe-2S faml
36	267	11.1	420	2 T31285	biphenyl dioxygena
37	251.5	10.4	374	2 AF0304	probable dioxygena
38	248	10.3	374	2 C85791	probable choline m
39	248	10.3	374	2 G90942	probable choline m
40	248	10.3	374	2 B64941	probable choline m
41	233.5	9.7	382	2 G70946	probable dioxygena
42	224	9.3	446	2 T14542	choline monooxygen
43	223	9.3	439	2 T09214	choline monooxygen
44	214.5	8.9	186	2 JC5354	2-nitrotoluene dio
45	186	7.7	426	2 T08550	choline monooxygen

#### ALIGNMENTS

RESULT 1  
JS0071  
naphthalene dioxygenase (EC 1.14.12.-) ndob protein - Pseudomonas putida  
C:Species: Pseudomonas putida  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999  
C:Accession: JS0071  
R:Kurkela, S.; Leivaeslahti, H.; Palva, E.T.; Teeri, T.H.  
Gene 73, 355-362, 1988  
A:Title: Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase and a characterizing protein of Pseudomonas putida  
A:Reference number: JS0070; MUID:89211973; PMID:3243438  
A:Accession: JS0071  
A:Molecule type: DNA  
A:Residues: 1-449 <KUR>  
A:Cross-references: GB:M23914; NID:9151392; PIDN:AAB47591.1; PID:9151394  
C:Comment: Naphthalene dioxygenase system is composed of three proteins.  
C:Genetics:  
A:Gene: ndob  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein F;1-119/Domain: Rieske [2Fe-2S] homology <RSK>  
E;81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 99.8% Score 2403; DB 2; Length 449;  
Best Local Similarity 99.8%; Pred. No. 5.2e-184  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MNVNNKIIVSESGLSQKHLIHGDELFQHELTIFARNMLFTHDSLIPAGDYVYAKMG	60
DB	1	MNVNNKIIVSESGLSQKHLIHGDELFQHELTIFARNMLFTHDSLIPAGDYVYAKMG	60
QY	61	IDEVYISRONDGSIRAFNLVNCRRHRTLVSEAGNKKGFVCSYHGWFSGNGLQSVPE	120
DB	61	IDEVYISRONDGSIRAFNLVNCRRHRTLVSEAGNKKGFVCSYHGWFSGNGLQSVPE	120
QY	121	KDLYGSILNKKCGLTKEVAHVSEFHEITGCFDEQAPLMDYGDAAWYLEPFRKSGGL	180
DB	121	KDLYGSILNKKCGLTKEVAHVSEFHEITGCFDEQAPLMDYGDAAWYLEPFRKSGGL	180
QY	181	ELVGPCKVYIKANMKAPAEVFGDHYVHWTHASSLSRSGESIFSSILAGNAALPPGAGI	240
DB	181	ELVGPCKVYIKANMKAPAEVFGDHYVHWTHASSLSRSGESIFSSILAGNAALPPGAGI	240
QY	241	QMTSKYSGMGVLMDSYGVSHADLPELMAFGAGAOERLNKETGVYRARIYRSHLNCIV	300
DB	241	QMTSKYSGMGVLMDSYGVSHADLPELMAFGAGAOERLNKETGVYRARIYRSHLNCIV	300
QY	301	FPNNSMLTCSGVFKVWNPIDANTTEVWTAYIVKMDPEDLKRLASVORTGPAEFWES	360
DB	301	FPNNSMLTCSGVFKVWNPIDANTTEVWTAYIVKMDPEDLKRLASVORTGPAEFWES	360
QY	361	DDNDNNETASONGKRYQSRSDLSNLGFGEDVYGDVAVPGVVKSAIETSRGFRAY	420
DB	361	DDNDNNETASONGKRYQSRSDLSNLGFGEDVYGDVAVPGVVKSAIETSRGFRAY	420

Db 361 DDNDMETSANGKKYQSRDLSNLGFGEDYGDVAYPGVVGKSAIGETSYRGFYRAY 420  
Oy 421 QAHVSSNMMAEFHASTTWHTELTCTDR 449  
Db 421 QAHVSSNMMAEFHASTTWHTELTCTDR 449

## RESULT 2

S27632  
naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large chain  
C:Species: Pseudomonas sp.  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S27632; B49343  
R:Denome, S.A.; Young, K.D.  
Submitted to the EMBL Data Library, February 1992  
A:Description: Cloning and molecular characterization of genes involved in metabolism of  
A:Reference number: S27631  
A:Accession: S27632

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <DEN>  
A:Cross-references: EMBL:M60405  
R:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
J. Bacteriol. 175, 6890-6901, 1993  
A:Title: Metabolism of dibenzothioophene and naphthalene in Pseudomonas strains: complete  
A:Reference number: A49343; MUID:94042852; PMID:822631  
A:Accession: B49343

A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 14-462 <DE2>

A:Cross-references: GB:M60405; NID:9151195; PID:AAA16125.1; PID:9294351

A:Experimental source: strain C18

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
F:84-112/Domain: Rieske [2Fe-2S] homology <RSK>  
F:94,96,114,117/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 99.8%; Score 2403; DB 2; Length 462;  
Best Local Similarity 99.8%; Pred. No. 5.4e-184;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MNYNNKILVSSGSLQKHLHGDELFQHELTIFARNWLFTHDSLIPAPDYYTAKMG 60  
Db 14 MNYNNKILVSSGSLQKHLHGDELFQHELTIFARNWLFTHDSLIPAPDYYTAKMG 73  
Oy 61 IDEVIVSRONGSIRAFVNCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
Db 74 IDEVIVSRONGSIRAFVNCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 133  
Oy 121 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 134 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 193  
Oy 181 ELVGPGRKVVYIKAMWKAENFVGDAHYGVGTHASSLSRSGESIFSSLAGNAALPREGAGL 240  
Db 194 ELVGPGRKVVYIKAMWKAENFVGDAHYGVGTHASSLSRSGESIFSSLAGNAALPREGAGL 253  
Oy 241 QMTSKYSGMGVLDGYSGVASADLVPPELMAFGAKOERLNKEIDGVARITRSHLNCYV 300  
Db 254 QMTSKYSGMGVLDGYSGVASADLVPPELMAFGAKOERLNKEIDGVARITRSHLNCYV 313  
Oy 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVKDMPEDLKRLRLADSVQRTGPGAFWES 360  
Db 314 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVKDMPEDLKRLRLADSVQRTGPGAFWES 373  
Oy 361 DDNDMETSANGKKYQSRDLSNLGFGEDYGDVAYPGVVGKSAIGETSYRGFYRAY 420  
Db 374 DDNDMETSANGKKYQSRDLSNLGFGEDYGDVAYPGVVGKSAIGETSYRGFYRAY 433  
Oy 421 QAHVSSNMMAEFHASTTWHTELTCTDR 449  
Db 434 QAHVSSNMMAEFHASTTWHTELTCTDR 462

## RESULT 3

JN0644

naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large cha

C:Species: Pseudomonas putida

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999

C:Accession: JN0644

R:Simon, M.J.; Osslund, T.D.; Saunders, R.; Ensley, B.D.; Suggs, S.; Harcourt, A.; Su

Gene 127, 31-37, 1993

A:Title: Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida st

A:Reference number: JN0640; MUID:93252277; PMID:8486285

A:Accession: JN0644

A:Molecule type: DNA

A:Residues: 1-449 <STW>

A:Cross-references: GB:M63949; NID:9151384; PID:AAA25902.1; PID:9151387

C:Comment: This protein is a member of naphthalene dioxygenase multicomponent enzyme

C:Genetics:

A:Gene: nahA

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [

F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 97.4%; Score 2345; DB 2; Length 449;  
Best Local Similarity 96.7%; Pred. No. 2.2e-179;  
Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Oy 1 MNYNNKILVSSGSLQKHLHGDELFQHELTIFARNWLFTHDSLIPAPDYYTAKMG 60  
Db 1 MNYNNKILVSSGSLQKHLHGDELFQHELTIFARNWLFTHDSLIPAPDYYTAKMG 60  
Oy 61 IDEVIVSRONGSIRAFVNCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
Db 61 IDEVIVSRONGSIRAFVNCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
Oy 121 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 121 KELYGESLNKKCLGKEVARVESFHGFTYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Oy 181 ELVGPGRKVVYIKAMWKAENFVGDAHYGVGTHASSLSRSGESIFSSLAGNAALPREGAGL 240  
Db 181 ELVGPGRKVVYIKAMWKAENFVGDAHYGVGTHASSLSRSGESIFSSLAGNAALPREGAGL 240  
Oy 241 QMTSKYSGMGVLDGYSGVASADLVPPELMAFGAKOERLNKEIDGVARITRSHLNCYV 300  
Db 241 QMTSKYSGMGVLDGYSGVASADLVPPELMAFGSGKQERLNKEIDGVARITRSHLNCYV 300  
Oy 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVKDMPEDLKRLRLADSVQRTGPGAFWES 360  
Db 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVKDMPEDLKRLRLADSVQRTGPGAFWES 360  
Oy 361 DDNDMETSANGKKYQSRDLSNLGFGEDYGDVAYPGVVGKSAIGETSYRGFYRAY 420  
Db 361 DDNDMETSANGKKYQSRDLSNLGFGEDYGDVAYPGVVGKSAIGETSYRGFYRAY 420  
Oy 421 QAHVSSNMMAEFHASTTWHTELTCTDR 449  
Db 421 QAHVSSNMMAEFHASTTWHTELTCTDR 449

## RESULT 4

C55217

polycyclic aromatic hydrocarbon dioxygenase (EC 1.14.12.-) iron-sulfur protein large

C:Species: Pseudomonas putida

C:Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 20-Jun-2000

C:Accession: C55217

R:Takizawa, N.; Kaida, N.; Toriige, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara

J. Bacteriol. 176, 2444-2449, 1994

A:Title: Identification and characterization of genes encoding polycyclic aromatic hy

A:Reference number: A55217; MUID:94209249; PMID:8157615

A:Accession: C55217

A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-449 <TAK>  
A:Cross-references: GB:AB004059; GB:D16629; NID:g2189972; PID:BAA20391.1; PID:g391844  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:71-119/Domains: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 95.1%; Score 2291; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 4.6e-175;  
Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

OY 1 MNVNNKIIVSESGLSQKKLHIGDEELFOHELKTIFARWMLFLTHDSLIPACEDYYTAMG 60  
||| ||||||||| :||| :|||||||  
Db 1 MNVNNKIIVSESGLSQKKLHIGDEELFOHELKAVRARMWLFLTHDSLIPSDDYATAMG 60  
  
OY 61 IDEVTVSRQNDGSIRAFPLNVCRRHKGTLYSVAGNAKFVCSYHGCGSGNGELOSVPE 120  
||||| ||||||| :||||| :|||||  
Db 61 IDEVTVSRQSDGSIRAFPLNVCRRHKGTLYVNAEAGNAKFVCSYHRWGSGNGELOSVPE 120  
  
OY 121 KDLYGESLNKKCLGLKEVARVESFHGFTYGCFFDOEAPILMDYLGDAAWLEPMFKHSGL 180  
||| :||||| :||||| :||||| :|||||  
Db 121 KELYESLNKKCLGLKEVARVESFHGFTYACIDOEAPILMDYLGDAAWLEPIFKHSGL 180  
  
OY 181 ELVGPGRGVVTKANKKAFAENPVGDAYHVGMTHASLSLGSISPSLGNALPREGAGL 240  
||| | ||||| :||| :||| :||| :|||  
Db 181 ELVGPGRGVVTKANKKAFAENPVGDAYHVGMTHASSLCTGESIFSSLGNVLPREGAGL 240  
  
OY 241 QMTSKYSGSMGLMDGYSGVSHADLVPELMAFGCAKQERLNKEIGDVARITYRSLNTCV 300  
||| ||||| :||| :||| :||| :|||  
Db 241 QMTSKYSGSMGLMDGYSGVSHADLVPELMAFGCAKQERLNKEIGDVARITYRSLNTCV 300  
  
OY 301 FPNNSMGLTCSCGYFKWNPIDANTTEVTWTYAIVEKDMPEDLKRLADSVOPTVGPAEWES 360  
||| :||||| :||||| :||||| :|||||  
Db 301 FPNNSVLTCSCGFYKWNPIDANTEVTWTYAIVEKMPEDLKRLLADVAORTFGPAGWES 360  
  
OY 361 DDNDMEFRASONGKRGYQRSDSLNLGFGPDYVDGDAYPCVYGKSALGETSYRCFYAX 420  
||| :||||| :||||| :||||| :|||||  
Db 361 DDNDMEFRASONGKRGYQRSDSLNLGFGKDVGDAYPGVYGKSALGETSYRCGYRAY 420  
  
OY 421 QAHVSSSNMAEFENHASTWHTELTKTTDR 449  
||| :||||| :||||| :|||||  
Db 421 QAHVSSSNMAEFEDAASSTWHTELTKTTDR 449

RESULT 5  
JC5352  
2-nitrotoluene dioxygenase (EC 1.14.12.-) iron-sulfur protein large chain - Pseudomonas  
N:Alternate names: 2-nitrotoluene iron-sulfur protein alpha chain  
C:Species: Pseudomonas sp.  
C>Date: 27-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 20-Aug-1999  
C:Accession: JC5352  
R:Parales, J.V.; Kumar, A.; Parales, R.E.; Gibson, D.T.  
A:Title: Cloning and sequencing of the genes encoding 2-nitrotoluene dioxygenase from *P.*  
A:Reference number: JC5350; MUID:97128768; PMID:8973308  
A:Accession: JC5352  
A:Molecule type: DNA  
A:Residues: 1-447 <PAR>  
A:Cross-references: GB:U49504; NID:g1773273; PID:AAB40383.1; PID:g1773277  
A:Experimental source: strain JS42  
C:Comment: This enzyme catalyzes the addition of dioxygen to the aromatic nucleus of 2-n  
A:Gene: ntdac  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:69-117/Domains: Rieske [2Fe-2S] homology <RSK>  
F:79,81,99,102/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status prede

Query Match 85.3%; Score 2055; DB 2; Length 447;  
Best Local Similarity 83.7%; Pred. No. 3.2e-156;  
Matches 376; Conservative 35; Mismatches 36; Indels 2; Gaps 1;

QY 1 MNVNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWMLFTJHDSLIPAPGDVYTAKMG 60  
 Db 1 MSYQN--LVSEAGLTQKLLIHGDKELFQHELKTIFARNWMLFTJHDSLIPSPDGYTAKMG 58  
 QY 61 IDEVIVSQNOSSTAFPLNVCNHRKTLTVSVDAAGNAKGFVCSYHWMGSGSNELQSVFFE 120  
 Db 59 VDEVIVSQNOSSTAFPLNVCNHRKTLTVHEAGNAKGFVCSYHWMGSGSNELQSVFFE 118  
 QY 121 KVLGESLNNKICGLKEVARVESFHFYTCGFDQAPPLMDYLGAAWYLEPMFKHSGGL 180  
 Db 119 KELYGDALIKKICGLKEVPRISEFHFYTCGCDADAPPLIDYLGAAWYLEPTFHHSGL 178  
 QY 181 ELVGPBGKVVYIKANKKAPAEENFGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPEGAGL 240  
 Db 179 ELVGPBGKVVYIKANKKAPAEENFGDIYHVGWTHAALRAGQSVFSSLAGNAKLPEEGAGL 238  
 QY 241 QMTSKYSGSMGVLMDCYSGSVHADYLPPELMAFGAKOERLNKEIDYARIRSHLNCTV 300  
 Db 239 QMTSKYSGSMGVLMDCYSGSVHADYLPPELMAFGAKOERLNKEIDYARIRSHLNCTV 298  
 QY 301 PNNNSMLTCSGFGFKWNPIDANTTEWYTAIYEKMPEDLKRRLDSVQRTYGPAGFMES 360  
 Db 299 PNNNSFLTGSATFKWNPIDENTTEWYTAIYEKMPEDLKRRLDAQAORSIGPAGFMES 358  
 QY 361 DDNDNMETASONGKKYKYSRSDLSNLGFGEDVYGDAYPYGVGKSAIGETSYRGFYRAY 420  
 Db 359 DDNDNMETLSQNAKKYKYSNSPDQIASLFGKQDVYGDYDECYPGVVGKSAIGETSYRGFYRAY 418  
 QY 421 QAHVSSNMWAEFEHASSTWHTELTKTDR 449  
 Db 419 QAHISSNMWAEFEHASSRWHTELTKTDR 447  
 RESULT 6  
 T31134  
 naphthalene dioxygenase (EC 1.14.12.-) large chain - Sphingomonas aromaticivorans  
 C.Species: Sphingomonas aromaticivorans  
 C.Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
 C.Accession: T31134  
 R.Romine, M.F.: Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.  
 submitted to the EMBL Data Library, July 1998  
 A.Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom  
 A.Reference number: Z20992  
 A.Accession: T31134  
 A.Status: preliminary; translated from GB/EMBL/DBJ  
 A.Molecule type: DNA  
 A.Residues: 1-459 <ROW>  
 A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378275; PIDN:AAD03858.1  
 C:Genetics:  
 A:Gene: bphA1f  
 A:Genome: plasmid pN11  
 C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [  
 C.Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
 C:80;82,100,103/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status  
 Query March 39.7%; Score 956; DB 2; Length 459;  
 Best Local Similarity 43.9%; Pred. No. 2e-68;  
 Matches 192; Conservative 71; Mismatches 162; Indels 12; Gaps 7;  
 QY 8 LVSESGLSQKHLIHGDEELFQHELKTIFARNWMLFTJHDSLIPAPGDVYTAKMGIDEVIVS 67  
 Db 7 LVDAANASQSRVFEWDQOVYQLELRIFRSCWMLMGHDSLVPKRGDFTTYMAEDRVLIS 66  
 QY 68 RQNDGSTAFPLNVCNHRKTLTVSVDAAGNAKGFVCSYHWMGSGSNELQSVFFPKDYSES 127  
 Db 67 RQPDSTLAFNINSCNHRNQICHDQSGSAKAFVCYHGMVGGQDSLVDYMERECYHSD 126  
 QY 128 LNNKICGLKEVARVESFHFYTCGFDQAPPLMDYLGAAWYLEPMFK-HSGGLVLCGP 186  
 Db 127 LDKSLTGLAPL-RVETTYGFIIFGCHDPEAPSLDEDYLGFCWYTLDTIMDGPBGGLLGP 185  
 QY 187 GKVVYIKANKKAPAEENFGDAYHVGWTHASSLSRSGESIFSSLAGNAAL-PPGAGLQMTSK 245

Db 186 LKSTLACNMKVPTEENVGDYGVHGTTHAALOMIGELAGLSGNRADPFDDLGIOFTWR 245  
QY YGSGWGLMDGSGVH--SADLVPELMAFGAKOERLNKEIGDVARIRSHLNCITVFN 303  
Db 246 HGRGFLIDNATAIHVARDGYKYLEETRGIRREKFGPE---KRLVGVHWNSTILPN 301  
QY NSMLTSCGVFWNIDANTTEWYTAIVKDMPEDLKRLRLADSVQRTVGAFWESDDN 363  
Db 302 CSFLTGWTGFKIMHPRGHEIEMWTYTWPKNADTETKRSIGREAIRSGTAGLESDDG 361  
QY 364 DNMEFASNGKKYQSRDSDLSNLFGEVDYGV--DAVYRGVYKSAIGESTYGFTRAYA 422  
Db 362 ENMSATVNNNGILTRKGRM--NSSMGKDREGPHVPGIVGVSFTGYSGYRFWQE 419  
QY 423 HVSSNMAEFHASTW 439  
Db 420 MLDAPDMAIRANDDTW 436

RESULT 7  
T31256  
terminal oxygenase component large chain homolog - Sphingomonas aromatiivorans plasmid  
C:Species: Sphingomonas aromatiivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31256  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A:Reference number: 220992  
A:Accession: T31256  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-450 <ROM>  
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378397; PIDN:AA03980.1  
C:Genetics:  
A:Genome: plasmid pNL1  
A:Note: bptala  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 32.1%; Score 773; DB 2; Length 450;  
Best Local Similarity 36.1%; Pred. No. 8, 2e-54;  
Matches 157; Conservative 88; Mismatches 184; Indels 6; Gaps 6;  
QY 1 MNNYKILVSSGSLQKHLIHDEELFOHELKTIFARNMLFTHDSLIPAPGDYTAAG 60  
Db 1 MNDSTADLVDSRTGQSSITASIEDIYQELERITGRCMLFLVHTSQIPKPDYRTFMG 60  
QY 61 IDEVIVSRONDGSTRAPLVNCHRGKTLVSEAGNAKGFVCSYHGWFSGNSGELSVPE 120  
Db 61 EDDVIVIRQKDGSIKAFINSCTHRGNRIRADRGNARAFCTYHGMFSFGALSGVPLE 120  
QY 121 KDLYGESLNKCKGLKEVAVESFHGFIYGCFOEAPPLMDYLGDAWYLEP-MKHHSG 179  
Db 121 NEAFVGEIDRTKFGILPTKVAEYKGLVFGCDWASPISLDVLGAKFELDWLAMPFG 180  
QY 180 LEVGPSPGVYIKAMKPAENFVGDAYHVGWTHASSLSRSGESI-FSSLA--GNAALPREG 237  
Db 181 SALLGETOKAMVLGTWKLPVENVCGDGYHLGNAHAGAAVAOSMDLTGLSVNSGVLDLG 240  
QY 238 AGLOMTSKYSGMGVLMGYSGVHSDLVPELMAFGAKOERLNKEIGDVARIR-ITYRSHL 296  
Db 241 -GLSVAGNMGHVLASLDGVSAYAFPPKPLLEYLEARNQVVIDRLGVRROQWGAQV 299  
QY 297 NCYTFPNNSMILTCGVEFWNPIDANTTEWYTAIVKDMPEDLKRLRLADSVQRTVGPA 356  
Db 300 NITTFPNQILPLGLNMFVYHPKGGQIEOMTWAMAENDMPAVALQILLENOCITFLFGLG 359  
QY 357 FWESDDNDNMTASONGKKYQSRDSDLSNLFGEVDYGVDAVYRGVYKSAIGESTYRGE 416  
Db 360 LFDNDGDNLACTEQSQKGMRTAQMDDVYTNMALGRSGKREG-FPGDIAAGLVSEHNRIF 418

QY 417 YRAYOAHVSSSNMAE 431  
Db 419 YRRQEHMAETWAE 433

RESULT 8  
A65031  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escheri  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: A65031; S49292  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.: Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: A65031  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <BIAT>  
A:Cross-references: GB:AE000340; GB:U00096; NID:g1788883; PIDN:AACT5591.1; PID:g17888  
A:Experimental source: strain K-12, substrain MG1655  
R:Tutlin, E.; Gasser, F.; Biville, F.  
submitted to the EMBL Data Library, September 1994  
A:Description: Cloning and sequencing of an E. coli gene homologous to dioxygenase of  
A:Reference number: S49292  
A:Accession: S49292  
A:Molecule type: DNA  
A:Residues: 1-19, 'A', 21-383, 'ATAPATANGVWK', 397-398, 451, 'R', 453, 'SAATATFALTLTIFQKPL  
A:Cross-references: EMBL:Z37966; NID:g550595; PIDN:CAA6018.1; PID:g550596  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:75-123/Domain: Rieske [2Fe-2S] homology <RSK>  
F:85,87,105,108/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 30.6%; Score 736; DB 2; Length 453;  
Best Local Similarity 36.4%; Pred. No. 7, 4e-51;  
Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;  
QY 20 IHGDELFQHELKTFANMLFTHDSLIPAPGVNAKMIDEVIVSRONDGSTRAPLN 79  
Db 24 IYTDPIYQLEERITFGKWLFLAHESQIPKPGFPNTYKEDAVVYVRKQDSIKAFLN 83  
QY 80 VCRHKGTLVSEAGNAKGFVCSYHGWFSGNSGELSVPEKDLYGESLNKCKGLKEVA 139  
Db 84 CCRHARAVSYADCGNTRAFCTCPYHWSYGINGELIDVLEPRAYPGCLCKSHWGLENVP 143  
QY 140 RVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMP-KHSGLELVGPGRVYIKAMKAP 198  
Db 144 CVESYKGLIFGNWDTASAGLRDYLGDIAWYLDGMIDRRGGTETIVGVQKWVINCMKFP 203  
QY 199 AENFVGDAYHVGWTHASSL-----RSGESIFSSLAGNAALP----- 234  
Db 204 AEOFAVDYHALFSAHSAVQVLGAKDGSCKRLD-----GGTARVWETAKDALQFG 256  
QY 235 --PEGAGLOMTSKYSGMGVLMGYSGVHSDLVPELMAFGAKOERLNKEIGDVARIR 292  
Db 257 QDGHSGFFETKEKPDANVW--DGAVSSYRETYAE-----AEORLGEVRLRL 303  
QY 293 RSHLCTVFPNNSMILTCGVEFWNPIDANTTEWYTAIVKDMPEDLKRLRLADSVQRTV 352  
Db 304 AGHNN--IFPTLSWNGATRLRVHPRGPDQVEWAFCTITKKAASDEKKAFFENSAFRAF 361  
QY 353 GPAGFWESDDNDNMTASONGKKYQSRDSDLSNLFGEVDYGVDAVYRGVYKSAISET 412  
Db 362 GPAGFLDQDSENWCEIOKLKLGHRARNSKLCLEHGLGQEKRRDDGIPGLT-NITSETA 420  
QY 413 YRGFYRAYOAHVSSSNMAEFHASTWHTLTK 445  
Db 421 ARGWYQNWADLLSSSEWQEVLDKTAAYQOEYWK 453

RESULT 9











NO

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:46:47 : Search time 29.1667 Seconds  
(without alignments)  
2051.298 Million cell updates/sec

Title: US-09-843-250-2

Perfect score: 2408  
Sequence: 1 MNVNMKIIIVSEGLSQKHLI.....AEFEHASSYWHTELTFTDR 449

Scoring table: BLOSUM62

Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2408	100.0	449	21	AAAB12565
2	2408	100.0	449	21	AAAB12566
3	2408	100.0	449	21	AAAB12567
4	2407	100.0	449	21	AAAB12582
5	2405	99.9	449	21	AAAB12581
6	2404	99.8	449	21	AAAB12579
7	2404	99.8	449	21	AAAB12580
8	2403	99.8	449	21	AAAB12568
9	2403	99.8	449	21	AAAB12577
10	2401	99.7	449	21	AAAB12578

11	2401	99.7	449	21	AAAB12583	Naphthalene dioxyg
12	2398	99.6	448	21	AAAB12584	Naphthalene dioxyg
13	2398	99.6	449	21	AAAB12569	NDO related comple
14	2350	97.6	449	21	AAAB12570	NDO related comple
15	2342	96.4	449	21	AAAB12571	NDO related comple
16	2296	95.3	449	21	AAAB12572	NDO related comple
17	2223	92.3	449	21	AAAB12573	NDO related comple
18	2193	91.1	447	21	AAAB12574	NDO related comple
19	2056	85.4	447	21	AAAB12575	NDO related comple
20	1972.5	81.9	451	21	AAAB12576	NDO related comple
21	1094	45.4	452	21	AAAB29474	Burkholderia sp. C
22	736	30.6	453	22	AAU34660	Benzene ring hydro
23	660.5	27.4	898	21	AAV83939	Terrabacter oxidas
24	657.5	27.3	443	23	AAO17350	Aromatic dihydrodi
25	622.5	25.9	459	16	AAE66729	Chimeric PCB decomp
26	617	25.6	458	21	AAV81990	Chimeric PCB decomp
27	614	25.5	458	21	AAV81989	Benzene dioxygenas
28	553.5	23.0	427	14	AAE32085	C glutamicum prote
29	376	15.6	490	22	AAE92386	Klebsiella pneumon
30	375.5	15.6	385	22	AAU36099	Corynebacterium 91
31	357	14.8	497	22	AAE79227	Pseudomonas aerugi
32	340.5	14.1	424	22	AAU33601	C album choline mo
33	239	9.9	435	22	AAE62634	C album choline mo
34	235	9.8	433	22	AAE62633	C album choline mo
35	235	9.8	433	22	AAE62635	Atirplex hortensis
36	233	9.7	438	23	AAE06798	Sugarbeet choline
37	224	9.3	446	19	AAE69563	Spinach choline mo
38	223	9.3	439	19	AAE69562	Human oxygenase-11
39	166	8.1	99	23	AAE35375	Herbicideally activ
40	166	7.7	426	23	AAE93163	Novel human diagno
41	183	7.6	1193	22	AAE18089	Rieske motif in 98
42	181	7.5	35	19	AAE79018	Rieske motif in 67
43	173	7.2	35	19	AAE79019	Rieske motif in DN
44	164	6.8	35	19	AAE79020	Oxidase amino acid
45	126	5.2	405	19	AAE80332	

ALIGNMENTS

RESULT 1	
AAAB12565	AAAB12565 standard; Protein: 449 AA.
ID	AAAB12565
AC	AAAB12565;
DT	09-NOV-2000 (first entry)
DE	Naphthalene dioxygenase mutant F352V protein sequence SEQ ID NO:2.
XX	
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KW	Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
OS	Pseudomonas sp.
OS	Synthetic.
XX	
PN	WO200037480-A1.
XX	
PD	29-JUN-2000.
XX	
PF	26-OCT-1999; 99WO-US25079.
XX	
PR	26-OCT-1998; 98US-0105575.
XX	
PA	(IOWA ) UNIV IOWA RES FOUND.
XX	
PI	Parales R, Gibson D, Resnick S, Lee K;
XX	
DR	WPI: 2000-452174/39.
DR	N-PSDB; AAA65339.
PT	Novel naphthalene dioxygenase mutant having a specific amino acid

```

RESULT 2
AAB12566
ID AAB12566 standard; Protein; 449 AA
XX
AC AAB12566;

```

QY	1	MNYNNKILVSESGISQXHLIHGDEELFOHELKITFARNMLFETHDSLIPAPGYVYAKKG	60
Db	1	MNYNNKILVSESGISQXHLIHGDEELFOHELKITFARNMLFETHDSLIPAPGYVYAKKG	60
QY	61	IDEVIVSRQNDGSLRAFLNVCRRHGKTLVSEVAGNKGVCVSHGMSGNGELOSPE	120
Db	61	IDEVIVSRQNDGSLRAFLNVCRRHGKTLVSEVAGNKGVCVSHGMSGNGELOSPE	120
QY	121	KDLGSESLNKKCLIEKLVARESEPHGFIYSCFOQEAPLMDYIGDAAWYLEPFIKHSGL	180

DB 121 KDLYGSESLNKKCLGLKEVARVESFHGFIYGCPEQDEAPRLMDYLGDAAWYLEPWFKHSGGL 180  
QY 181 ELVGPFGKVVIRANKMAPENFYGDAYHYGWTTHASSLRGSESTFSSLAGNAALPPGAGL 240  
DB 181 ELVGPFGKVVIRANKMAPENFYGDAYHYGWTTHASSLRGSESTFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVHSADLVEPELMAFGAKOERLNKEIGVRRARIYRSHLNCIV 300  
DB 241 QMTSKYSGMGVLMDSYGVHSADLVEPELMAFGAKOERLNKEIGVRRARIYRSHLNCIV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVEKMPEDLKRRLADSVORTVGPAGFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVEKMPEDLKRRLADSVORTVGPAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSDBSDLSNLGFGEDYGGDAVYPGVGKSAIGETSYRGYRAY 420  
DB 361 DDNDNMETASQNGKKYQSDBSDLSNLGFGEDYGGDAVYPGVGKSAIGETSYRGYRAY 420  
QY 421 QAHVSSSNMAEFHASTWHTLTCTTDR 449  
DB 421 QAHVSSSNMAEFHASTWHTLTCTTDR 449  
RESULT 3  
AAB12567  
ID AAB12567 standard; Protein: 449 AA.  
XX AAB12567;  
AC  
XX  
XX  
DT 09-NOV-2000 (first entry)  
DE NDO related complex alpha subunit protein sequence SEQ ID NO:15.  
XX  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
XX Pseudomonas sp.  
OS Synthetic.  
XX  
XX WO200037480-A1.  
PN  
XX  
XX 29-JUN-2000.  
PD  
XX 26-OCT-1999; 99WO-US25079.  
PF  
XX 26-OCT-1998; 98US-0105575.  
PR  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI: 2000-452174/39.  
DR N-PSDB; AAA65341.  
XX  
XX Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT .  
PS  
XX  
XX Claim 13; Page 100-102; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which

CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
SQ  
XX  
XX Sequence 449 AA:  
Query Match 100.0%; Score 2408; DB 21; Length 449;  
Best Local Similarity 100.0%; Pred. No. 4e-225;  
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNNYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYVAKMG 60  
DB 1 MNNYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYVAKMG 60  
QY 61 IDEVIVSRQNDGSIKRAFLNVCRRHGGKTLVSVKGNKAGFYCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIVSRQNDGSIKRAFLNVCRRHGGKTLVSVKGNKAGFYCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSESLNKKCLGLKEVARVESFHGFIYGCPEQDEAPRLMDYLGDAAWYLEPWFKHSGGL 180  
DB 121 KDLYGSESLNKKCLGLKEVARVESFHGFIYGCPEQDEAPRLMDYLGDAAWYLEPWFKHSGGL 180  
QY 181 ELVGPFGKVVIRANKMAPENFYGDAYHYGWTTHASSLRGSESTFSSLAGNAALPPGAGL 240  
DB 181 ELVGPFGKVVIRANKMAPENFYGDAYHYGWTTHASSLRGSESTFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVHSADLVEPELMAFGAKOERLNKEIGVRRARIYRSHLNCIV 300  
DB 241 QMTSKYSGMGVLMDSYGVHSADLVEPELMAFGAKOERLNKEIGVRRARIYRSHLNCIV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVEKMPEDLKRRLADSVORTVGPAGFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTYAIVEKMPEDLKRRLADSVORTVGPAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSDBSDLSNLGFGEDYGGDAVYPGVGKSAIGETSYRGYRAY 420  
DB 361 DDNDNMETASQNGKKYQSDBSDLSNLGFGEDYGGDAVYPGVGKSAIGETSYRGYRAY 420  
QY 421 QAHVSSSNMAEFHASTWHTLTCTTDR 449  
DB 421 QAHVSSSNMAEFHASTWHTLTCTTDR 449  
RESULT 4  
AAB12582  
ID AAB12582 standard; Protein: 449 AA.  
XX  
XX AAB12582;  
AC  
XX  
XX 09-NOV-2000 (first entry)  
DT  
XX  
XX Naphthalene dioxygenase mutant F352I protein sequence SEQ ID NO:36.  
DE  
XX  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
XX Pseudomonas sp.  
OS Synthetic.  
XX  
XX WO200037480-A1.  
PN  
XX  
XX 29-JUN-2000.  
PD  
XX 26-OCT-1999; 99WO-US25079.  
PF

XX 26-OCT-1998; 98US-0105575.  
PR  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI; 2000-452174/39.  
DR  
DR N-PSDB; AAA65356.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
XX  
PS Claim 9; Page 133-134; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SQ  
Query Match 100.0%; Score 2407; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 5e-225;  
Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNVNNKILVSSGSLQKHLHGDELPQHEKLTIFARWMLFLTHSLRPAPDDYTKMG 60  
DB 1 MNVNNKILVSSGSLQKHLHGDELPQHEKLTIFARWMLFLTHSLRPAPDDYTKMG 60  
QY 61 IDEVIVSRONGDSTRAPLNVCRHKGKTLVSVEAGNAKGFVCSYHMGFGSNGELQVPEE 120  
DB 61 IDEVIVSRONGDSTRAPLNVCRHKGKTLVSVEAGNAKGFVCSYHMGFGSNGELQVPEE 120  
QY 121 KDLYGESLNNKKCLCLKEYARVESFHGFTYGCFFDAPPLMDYLGDAAVLEPMKHSGL 180  
DB 121 KDLYGESLNNKKCLCLKEYARVESFHGFTYGCFFDAPPLMDYLGDAAVLEPMKHSGL 180  
QY 181 ELVGPPEKVVYKAMWKAPEAFNVGDVAVHGMTHASSLSRSGSTSSLAGNALPREGAGL 240  
DB 181 ELVGPPEKVVYKAMWKAPEAFNVGDVAVHGMTHASSLSRSGSTSSLAGNALPREGAGL 240  
QY 241 QMTSKYSGMGVLMDGYSGVSHADVLPELMAFGAKOERLKEIGDVARARYRSHLNTCV 300  
DB 241 QMTSKYSGMGVLMDGYSGVSHADVLPELMAFGAKOERLKEIGDVARARYRSHLNTCV 300  
QY 301 FPNNSMLTSCGVEFKVWNPIDANTTEVMTYAIVEKDMPEDLKRLADSVQRTIGPAGFWS 360  
DB 301 FPNNSMLTSCGVEFKVWNPIDANTTEVMTYAIVEKDMPEDLKRLADSVQRTIGPAGFWS 360  
QY 361 DDNDNMTASONGKKYOSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMTASONGKKYOSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420

QY 421 QAHVSSNMWAEFEHASTWHTLETKTDR 449  
DB 421 QAHVSSNMWAEFEHASTWHTLETKTDR 449  
RESULT 5  
AAB12581  
ID AAB12581 standard; Protein: 449 AA.  
XX  
XX AAB12581;  
AC  
XX 09-NOV-2000 (first entry)  
DE  
XX Naphthalene dioxygenase mutant F352L protein sequence SEQ ID NO:35.  
XX  
XX  
XX Pseudomonas sp. strain NCTB 9816-4; naphthalene dioxygenase: NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
XX Pseudomonas sp.  
OS Synthetic.  
XX  
XX WO200037480-A1.  
PN  
XX 29-JUN-2000.  
PD  
XX 26-OCT-1999; 99WO-US25079.  
PF  
XX 26-OCT-1998; 98US-0105575.  
PR  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI; 2000-452174/39.  
DR  
DR N-PSDB; AAA65355.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
XX  
PS Claim 9; Page 131-133; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SQ  
Query Match 99.98%; Score 2405; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 7.8e-225;  
Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWMLFTHDSLIPAGDYVYAKMG 60  
DB 1 MNYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWMLFTHDSLIPAGDYVYAKMG 60  
QY 61 IDEVYISRQNDGSIIRAFNLVNCGRHRTLVSEVGNNAKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVYISRQNDGSIIRAFNLVNCGRHRTLVSEVGNNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSELNKKCLGKLEVARVESFHGFIYGCFOEARPLMDYIGDAWYLEPMPFKHSGGL 180  
DB 121 KDLYGSELNKKCLGKLEVARVESFHGFIYGCFOEARPLMDYIGDAWYLEPMPFKHSGGL 180  
QY 181 ELVGPPEKVVYIKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPREGAGL 240  
DB 181 ELVGPPEKVVYIKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPREGAGL 240  
QY 241 QMTSKYSGMGVLMGYSGVHSDVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMGYSGVHSDVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVTAIVEKMDPELKRRLADSVORTVPAGFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVTAIVEKMDPELKRRLADSVORTVPAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGDVYGVGKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGDVYGVGKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449  
DB 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

RESULT 6  
AAB12579  
ID AAB12579 standard; Protein: 449 AA.  
XX  
AC AAB12579;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352A protein sequence SEQ ID NO:33.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
XX  
OS Synthetic.  
XX  
PN MO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI; 2000-452174/39.  
XX  
DR N-PSDB; AAA65353.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
substitution for preparing chiral diols for use in the polymer, resin,  
pharmaceutical or rubber industry and for carrying out bioremediation  
XX  
PS Claim 9; page 128-129; 151pp; English.  
XX

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(4,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzotran,  
CC dibenzochlorophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydro compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

Sequence 449 AA:

Query Match 99.8%; Score 2404; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 9.8e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWMLFTHDSLIPAGDYVYAKMG 60  
DB 1 MNYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWMLFTHDSLIPAGDYVYAKMG 60  
QY 61 IDEVYISRQNDGSIIRAFNLVNCGRHRTLVSEVGNNAKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVYISRQNDGSIIRAFNLVNCGRHRTLVSEVGNNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSELNKKCLGKLEVARVESFHGFIYGCFOEARPLMDYIGDAWYLEPMPFKHSGGL 180  
DB 121 KDLYGSELNKKCLGKLEVARVESFHGFIYGCFOEARPLMDYIGDAWYLEPMPFKHSGGL 180  
QY 181 ELVGPPEKVVYIKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPREGAGL 240  
DB 181 ELVGPPEKVVYIKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPREGAGL 240  
QY 241 QMTSKYSGMGVLMGYSGVHSDVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMGYSGVHSDVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVTAIVEKMDPELKRRLADSVORTVPAGFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVTAIVEKMDPELKRRLADSVORTVPAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGDVYGVGKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGDVYGVGKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449  
DB 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

RESULT 7  
AAB12580  
ID AAB12580 standard; Protein: 449 AA.  
XX  
AC AAB12580;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352T protein sequence SEQ ID NO:34.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
XX

KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
OS Pseudomonas sp.  
OS Synthetic.  
XX WO200037480-A1.  
XX  
XX 29-JUN-2000.  
XX  
XX 26-OCT-1999; 99WO-US25079.  
XX  
XX 26-OCT-1998; 98US-0105575.  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
XX  
XX  
XX Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI: 2000-452174/39.  
XX  
XX N-PSDB: AAA65354.  
XX  
XX Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
PS Claim 9; Page 129-131; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzoocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SO  
Query Match 99.8%; Score 2404; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 9.8e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 QMTSKYSGMGLMDGYSGVSHADLVPPELMAFGAKOERLNKEIGDYRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGLMDGYSGVSHADLVPPELMAFGAKOERLNKEIGDYRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCGSGYFKVWNPIDANTTEWWTYAIVEKMPEDLKRRLADSVORTGAPGWES 360  
DB 301 FPNNSMLTCGSGYFKVWNPIDANTTEWWTYAIVEKMPEDLKRRLADSVORTGAPGWES 360  
QY 361 DDNDNMETASQNGKKYQSRDSDLSNLGFGEDYGDAYPGVWKSALGERSYGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSRDSDLSNLGFGEDYGDAYPGVWKSALGERSYGFYRAY 420  
QY 421 QAHVSSNMAEFEEHASTWHELETKTDR 449  
DB 421 QAHVSSNMAEFEEHASTWHELETKTDR 449  
RESULT 8  
AAB12568  
ID AAB12568 standard; Protein: 449 AA.  
XX  
XX AAB12568;  
AC  
XX  
XX 09-NOV-2000 (first entry)  
XX  
XX  
XX NDO related complex alpha subunit protein sequence SEQ ID NO:16.  
XX  
XX Pseudomonas sp. strain NCIB 9616-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
XX Pseudomonas sp.  
OS Synthetic.  
OS  
XX WO200037480-A1.  
PN  
XX 29-JUN-2000.  
PD  
XX 26-OCT-1999; 99WO-US25079.  
XX  
XX 26-OCT-1998; 98US-0105575.  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
XX Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI: 2000-452174/39.  
XX  
XX N-PSDB: AAA65342.  
DR  
XX  
XX Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
PS Claim 13; Page 102-103; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzoocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a

CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.8%; Score 2403; DB 21; Length 449;

Best Local Similarity 99.8%; Pred. No. 1.2e-224;

Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNNKILVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPGDYVATKMG 60  
1 MNYNNKILVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPGDYVATKMG 60  
Db 61 IDEVIYSRQNDGSIKRAFLNVCRIHGRKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 61 IDEVIYSRQNDGSIKRAFLNVCRIHGRKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
Db 61 IDEVIYSRQNDGSIKRAFLNVCRIHGRKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGESLNNKCKLGLEKVARVESFHGFIYGCPOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
121 KDLYGESLNNKCKLGLEKVARVESFHGFIYGCPOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 121 KDLYGESLNNKCKLGLEKVARVESFHGFIYGCPOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPPEKVVYIKANKKAPAEVFGDAYHVGMTTHASSLRSGESIFSSLAGNAALPPREGAGL 240  
181 ELVGPPEKVVYIKANKKAPAEVFGDAYHVGMTTHASSLRSGESIFSSLAGNAALPPREGAGL 240  
Db 181 ELVGPPEKVVYIKANKKAPAEVFGDAYHVGMTTHASSLRSGESIFSSLAGNAALPPREGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVHSADLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
241 QMTSKYSGMGVLMDSYGVHSADLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
Db 241 QMTSKYSGMGVLMDSYGVHSADLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVEKVVNPIDANTTEVWTAIVEKMDPEDLKRLADSVORTVPAGFWES 360  
301 FPNNSMLTCSGVEKVVNPIDANTTEVWTAIVEKMDPEDLKRLADSVORTVPAGFWES 360  
Db 301 FPNNSMLTCSGVEKVVNPIDANTTEVWTAIVEKMDPEDLKRLADSVORTVPAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGAVPVGVKSAIGETSYRGFYRAY 420  
361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGAVPVGVKSAIGETSYRGFYRAY 420  
Db 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGAVPVGVKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNNAAEFHASTWTELTCTTDR 449  
421 QAHVSSNNAAEFHASTWTELTCTTDR 449  
Db 421 QAHVSSNNAAEFHASTWTELTCTTDR 449

RESULT 9

AAB12577  
ID AAB12577 standard; Protein; 449 AA.

XX AAB12577:

DT 09-NOV-2000 (first entry)

XX Naphthalene dioxygenase alpha subunit protein sequence SEQ ID NO:26.

XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;

XX inducible multi-component enzyme; alpha subunit; mutant; chiral diol;

XX polymer; resin; pharmaceutical; rubber industry; bioremediation.

XX Pseudomonas sp.

XX WO200037480-A1.

XX 29-JUN-2000.

XX 26-OCT-1999; 99MO-US25079.

XX 26-OCT-1998; 98US-0105575.

XX (IOWA ) UNIV IOWA RES FOUND.

XX Paraes R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

XX N-PSDB: AAA65351.

PT Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,

PT pharmaceutical or rubber industry and for carrying out bioremediation

XX Disclosure: Page 118-120; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzochiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC naphthalene dioxygenase (NDO) from the Pseudomonas sp. strain NCIB  
CC 9816-4, which is used in the exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.8%; Score 2403; DB 21; Length 449;

Best Local Similarity 99.8%; Pred. No. 1.2e-224;

Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNNKILVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPGDYVATKMG 60  
1 MNYNNKILVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPGDYVATKMG 60  
Db 1 MNYNNKILVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPGDYVATKMG 60  
QY 61 IDEVIYSRQNDGSIKRAFLNVCRIHGRKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
61 IDEVIYSRQNDGSIKRAFLNVCRIHGRKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
Db 61 IDEVIYSRQNDGSIKRAFLNVCRIHGRKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGESLNNKCKLGLEKVARVESFHGFIYGCPOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
121 KDLYGESLNNKCKLGLEKVARVESFHGFIYGCPOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 121 KDLYGESLNNKCKLGLEKVARVESFHGFIYGCPOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPPEKVVYIKANKKAPAEVFGDAYHVGMTTHASSLRSGESIFSSLAGNAALPPREGAGL 240  
181 ELVGPPEKVVYIKANKKAPAEVFGDAYHVGMTTHASSLRSGESIFSSLAGNAALPPREGAGL 240  
Db 181 ELVGPPEKVVYIKANKKAPAEVFGDAYHVGMTTHASSLRSGESIFSSLAGNAALPPREGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVHSADLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
241 QMTSKYSGMGVLMDSYGVHSADLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
Db 241 QMTSKYSGMGVLMDSYGVHSADLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVEKVVNPIDANTTEVWTAIVEKMDPEDLKRLADSVORTVPAGFWES 360  
301 FPNNSMLTCSGVEKVVNPIDANTTEVWTAIVEKMDPEDLKRLADSVORTVPAGFWES 360  
Db 301 FPNNSMLTCSGVEKVVNPIDANTTEVWTAIVEKMDPEDLKRLADSVORTVPAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGAVPVGVKSAIGETSYRGFYRAY 420  
361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGAVPVGVKSAIGETSYRGFYRAY 420  
Db 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYVGAVPVGVKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNNAAEFHASTWTELTCTTDR 449  
421 QAHVSSNNAAEFHASTWTELTCTTDR 449  
Db 421 QAHVSSNNAAEFHASTWTELTCTTDR 449

RESULT 10

AAB12578  
ID AAB12578 standard; Protein: 449 AA.  
XX  
AC AAB12578;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352G protein sequence SEQ ID NO:32.  
XX  
KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
XX polymer; resin; pharmaceutical; rubber industry; bioremediation.  
OS Pseudomonas sp.  
XX Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
DR N-PSDB: AAA65352.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Claim 9; Page 126-127; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzofluorene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 449 AA;  
XX  
Query Match 99.7%; Score 2401; DB 221; Length 449;  
Best Local Similarity 99.8%; Pred. NO. 1.9e-22;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 61 IDEVIVSRONDGSIIRAFILNVCRRHGRKTLVSVAGNAKGFVCSYHMGFGSGNGLQSVPE 120  
QY 121 KDLYGSESLNKKKCLGLKEVARVSEFGFTYCGFDQCAPPLMDYLGDAAAYLEPMKRHSGGL 180  
DB 121 KDLYGSESLNKKKCLGLKEVARVSEFGFTYCGFDQCAPPLMDYLGDAAAYLEPMKRHSGGL 180  
QY 181 ELVGPFGKVVIRANKKAPAEENFVGDAHYHGWTHASLSRSGESIFSSLAGNAALPPEGAGL 240  
DB 181 ELVGPFGKVVIRANKKAPAEENFVGDAHYHGWTHASLSRSGESIFSSLAGNAALPPEGAGL 240  
QY 241 QMTSKYSGSMGVLMDCYSGVSHADLVPELMAFGAKOERLKEIGDVARIRYRSHLNCVY 300  
DB 241 QMTSKYSGSMGVLMDCYSGVSHADLVPELMAFGAKOERLKEIGDVARIRYRSHLNCVY 300  
QY 301 FPNNSMLTCSGVKFWNPIDANTTEVWYIAVEKMPEDLKRRLADSVQRVPGAGFWES 360  
DB 301 FPNNSMLTCSGVKFWNPIDANTTEVWYIAVEKMPEDLKRRLADSVQRVPGAGFWES 360  
QY 361 DDNDNMETASQNGKKYQSRSDLSNLGFGEDVYGDAYPGVGSALGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSRSDLSNLGFGEDVYGDAYPGVGSALGETSYRGFYRAY 420  
QY 421 QAHVSSSNNAEFEEHASSTWHTELTYTDR 449  
DB 421 QAHVSSSNNAEFEEHASSTWHTELTYTDR 449  
RESULT 11  
AAB12583  
ID AAB12583 standard; Protein: 449 AA.  
XX  
AC AAB12583;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352W protein sequence SEQ ID NO:58.  
XX  
KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
DR N-PSDB: AAA65376.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Example 7; Page 142-144; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which



CC involves contacting them with naphthalene, biphenyl, phenanthrene, and  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthantrone or 3,4-dihydroxy-3,4-dihydronaphthantrone which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothioephene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 449 AA;

Query Match 99.7%; Score 2401; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 1.9e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

	Query Match	99.7%	Score 2401	DB 21	Length 449
	Best Local Similarity	99.8%	Pred. No. 1.9e-224		
	Matches 448	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1 MNYNNKILVSSGSLQKRLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYATAKG	60			
Db	1 MNYNNKILVSSGSLQKRLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYATAKG	60			
QY	61 IDEVIVSRQNDGSIIRAFNLVNCRRHGRKTLVSVEAGNAKGFVCSYHGWGFGSGNELQSVPE	120			
Db	61 IDEVIVSRQNDGSIIRAFNLVNCRRHGRKTLVSVEAGNAKGFVCSYHGWGFGSGNELQSVPE	120			
QY	121 KDLIGESLNKKCLGLKEVARSEFHGFITYGCFDQAPPLMDYLGAANYLEPMFKHSGGL	180			
Db	121 KDLIGESLNKKCLGLKEVARSEFHGFITYGCFDQAPPLMDYLGAANYLEPMFKHSGGL	180			
QY	181 ELVAPPGKVYIKANMKKAPAEVNFVDGAYHVGMTHASSLSGSEIFSSLAGNALPPEGAGL	240			
Db	181 ELVAPPGKVYIKANMKKAPAEVNFVDGAYHVGMTHASSLSGSEIFSSLAGNALPPEGAGL	240			
QY	241 QMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLNKEIDVRRARYRSHLNCY	300			
Db	241 QMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLNKEIDVRRARYRSHLNCY	300			
QY	301 FPNNSMLTCSGVFKVMPIDANTTEVMYTAIVKDMPEDLKRLRADSQRTVGPAGFWS	360			
Db	301 FPNNSMLTCSGVFKVMPIDANTTEVMYTAIVKDMPEDLKRLRADSQRTVGPAGFWS	360			
QY	361 DDNDNMETASONGKKYOSRDSLLSNLFCGEDVYGDVAYPGVCGKSAIGETSYRGCFYAY	420			
Db	361 DDNDNMETASONGKKYOSRDSLLSNLFCGEDVYGDVAYPGVCGKSAIGETSYRGCFYAY	420			
QY	421 QAHVSSSNMAFEFHAASSTWHTELTKTDR	449			
Db	421 QAHVSSSNMAFEFHAASSTWHTELTKTDR	449			
RESULT 12					
ID	AAB12584	standard; Protein; 448 AA.			
XX					
AC	AAB12584				
XX					
DT	09-NOV-2000	(first entry)			
XX					
DE	Naphthalene dioxygenase mutant F352Y protein sequence SEQ ID NO:59.				
XX					
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;				
KM	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;				
XX	polymer; resin; pharmaceutical; rubber industry; bioremediation.				
OS	Pseudomonas sp.				
OS	Synthetic.				
XX					
XX	WO200037480-A1.				

29-JUN-2000.  
26-OCT-1999; 99WO-US25079.  
26-OCT-1998; 98US-0105575.  
(IOWA ) UNIV IOWA RES FOUND.  
Parales R, Gibson D, Resnick S, Lee K;  
WPI: 2000-452174/39.  
N-PSDB; AAA65377.  
Novel naphthalene dioxygenase mutant having a specific amino acid substitution for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry and for carrying out bioremediation

Example 7; Page 144-145; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-) (1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene 1,2-dihydrodiol, cis-1,2-dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzo[c]chloro[1,1]-ene, anthracene, phenanthrene, dibenzo[1,4]dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothioephene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Sequence 448 AA;

Very Match	99.63;	Score 2398;	DB 21;	length 448;
1st Local Similarity	99.88;	Pred. No. 3.7e-224;		
Mismatches 447; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Qy	1	MNNKKILVYSSGSLQKRLHGDDEEPOHEKLTLEARNMYLFTHDSLIPAPGVYPAKKMG	60
Dd	1	MNYNNKILVYSSGSLQKRLHGDDEEPOHEKLTLEARNMYLFTHDSLIPAPGVYPAKKMG	60
Qy	61	IDEVIVSRQNDGSIIRAFILNVCRRHGKTLVSEAGNAGKFGVSYHGMGFGSNGELQSVPE	120
Dd	61	IDEVIVSRQNDGSIIRAFILNVCRRHGKTLVSEAGNAGKFGVSYHGMGFGSNGELQSVPE	120
Qy	121	KDLVGSILNKKKCLGLKEVARYESHGFYYGCFDEAPRLMDYLGDAANYLEPMFKHSGGL	180
Dd	121	KDLVGSILNKKKCLGLKEVARYESHGFYYGCFDEAPRLMDYLGDAANYLEPMFKHSGGL	180
Qy	181	ELVYGPCKVVIKAMMKAPAEVNFVDAHVGVTHASSLSRGSSEIFSSLAGNAALPPEGAGL	240
Dd	181	ELVYGPCKVVIKAMMKAPAEVNFVDAHVGVTHASSLSRGSSEIFSSLAGNAALPPEGAGL	240
Qy	241	QMTSKYSGGMVLMDDGYSGVSHADLVPELMAFGAKKBERLNKEIGDVRATYRSHNLCTY	300
Dd	241	QMTSKYSGGMVLMDDGYSGVSHADLVPELMAFGAKKBERLNKEIGDVRATYRSHNLCTY	300
Qy	301	FPNNSMLTCSGVFVYVMPIDANTTEVWTVAIVEKQMDPEDLKRLRLASVQRTVPAPGFWES	360
Dd	301	FPNNSMLTCSGVFVYVMPIDANTTEVWTVAIVEKQMDPEDLKRLRLASVQRTVPAPGFWES	360



PT -  
XX  
PS  
XX  
XX  
Claim 13; Page 105-107; 151pp; English.  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (1) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX  
SQ Sequence 449 AA:  
Query Match 97.6%; Score 2350; DB 21; Length 449;  
Best Local Similarity 96.9%; Pred. No. 1.7e-219;  
Matches 435; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MNYNKKLVSESGLSQKHLLHGDEELFOHELKTFARNMWLFTHDSLTPAGDYVYAKMG 60  
DB 1 MNYNKKLVSESGLSQKHLLHGDEELFOHELKTFARNMWLFTHDSLTPAGDYVYAKMG 60  
QY 61 IDEVIYSRQNDGSIKRAFLNVCRRHGGKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIYSRQNDGSIKRAFLNVCRRHGGKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSELNKKCGLGKLVANVESHGCTTCFQDEAPRLMDYLGDAWYLEPMFKHSGGL 180  
DB 121 KDLYGSELNKKCGLGKLVANVESHGCTTCFQDEAPRLMDYLGDAWYLEPMFKHSGGL 180  
QY 121 KELYGEELNKKCGLGKLVANVESHGCTTCFQDEAPRLMDYLGDAWYLEPMFKHSGGL 180  
DB 121 KELYGEELNKKCGLGKLVANVESHGCTTCFQDEAPRLMDYLGDAWYLEPMFKHSGGL 180  
QY 181 ELVGPPEKVVYIKAMKRAPNFVGDAYHVGWTHASSLRSESTFSSLAGNAALPPREGAGL 240  
DB 181 ELVGPPEKVVYIKAMKRAPNFVGDAYHVGWTHASSLRSESTFSSLAGNAALPPREGAGL 240  
QY 241 QMTSKYSGMGLMDGYSGVHSAADLVPELMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGLMDGYSGVHSAADLVPELMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300  
QY 301 FPNNSMITCGGVKRVWNPIDANTTEWTVYAIIVEKMPEDLKRRLADSVORTVPAGFWES 360  
DB 301 FPNNSMITCGGVKRVWNPIDANTTEWTVYAIIVEKMPEDLKRRLADSVORTVPAGFWES 360  
QY 361 DDNDNMFTASONGKKYOSRSDLSNLGFGEDYGDVAVGVYGGKSAIGTSYRGFEFRAY 420  
DB 361 DDNDNMFTASONGKKYOSRSDLSNLGFGEDYGDVAVGVYGGKSAIGTSYRGFEFRAY 420  
QY 421 QAHVSSSNMAEFHASSTWHTTELKTTDR 449  
DB 421 QAHVSSSNMAEFHASSTWHTTELKTTDR 449  
RESULT 15  
AAB12571  
ID AAB12571 standard; Protein; 449 AA.  
XX  
AC AAB12571;  
XX  
DT 09-NOV-2000 (first entry)

XX  
DE NDO related complex alpha subunit protein sequence SEQ ID NO:19.  
XX  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
XX  
OS Synthetic.  
FH Key Location/Qualifiers  
FT Misc-difference 35  
FT /note= "unspecified"  
PN MO200037480-A1.  
PN  
PD 29-JUN-2000.  
XX  
XX 26-OCT-1999; 99MO-US25079.  
PF  
XX 26-OCT-1998; 98US-0105575.  
PR  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
XX  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
DR N-PSDB; AAA65345.  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
PS  
XX  
XX  
XX  
PS Claim 13; Page 107-109; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (1) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX  
SQ Sequence 449 AA:  
Query Match 96.4%; Score 2322; DB 21; Length 449;  
Best Local Similarity 95.8%; Pred. No. 9.2e-217;  
Matches 430; Conservative 12; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MNYNKKLVSESGLSQKHLLHGDEELFOHELKTFARNMWLFTHDSLTPAGDYVYAKMG 60  
DB 1 MNYNKKLVSESGLSQKHLLHGDEELFOHELKTFARNMWLFTHDSLTPAGDYVYAKMG 60  
QY 61 IDEVIYSRQNDGSIKRAFLNVCRRHGGKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIYSRQNDGSIKRAFLNVCRRHGGKTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 09:02:32 ; Search time 7 Seconds  
(without alignments)  
1294.310 Million cell updates/sec

Title: US-09-843-250-32  
Perfect score: 2410  
Sequence: 1 MNVNNKILVSEGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEM\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCr\_NEM\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEM\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEM\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCrUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEM\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEM\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEM\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	734	30.5	453	10	US-09-815-242-10253
2	377.5	15.7	385	10	US-09-815-242-11692
3	376	15.6	490	9	US-09-738-626-6140
4	333.5	13.8	424	10	US-09-815-242-5097
5	185	7.7	35	10	US-09-776-490-19
6	185	7.7	35	10	US-09-776-491-19
7	174	7.2	35	10	US-09-776-490-20
8	174	7.2	35	10	US-09-776-491-20
9	168	7.0	35	10	US-09-776-490-21
10	168	7.0	35	10	US-09-776-491-21
11	118.5	4.9	354	1	US-08-976-0630-4
12	110	4.6	35	10	US-09-776-490-12
13	110	4.6	35	10	US-09-776-491-12
14	109	4.5	35	10	US-09-776-490-14
15	109	4.5	35	10	US-09-776-491-14
16	108	4.5	35	10	US-09-776-490-13
17	108	4.5	35	10	US-09-776-490-15
18	108	4.5	35	10	US-09-776-491-13
19	108	4.5	35	10	US-09-776-491-15

20	107.5	4.5	548	9	US-10-047-542-78	Sequence 78, Appl
21	106	4.4	35	10	US-09-776-490-9	Sequence 9, Appl
22	106	4.4	35	10	US-09-776-491-9	Sequence 9, Appl
23	104.5	4.3	951	9	US-09-924-097-15	Sequence 15, Appl
24	104	4.3	35	10	US-09-776-490-18	Sequence 18, Appl
25	104	4.3	35	10	US-09-776-491-18	Sequence 18, Appl
26	103	4.3	35	10	US-09-776-490-11	Sequence 11, Appl
27	103	4.3	35	10	US-09-776-491-11	Sequence 11, Appl
28	102	4.2	35	10	US-09-776-490-8	Sequence 8, Appl
29	102	4.2	35	10	US-09-776-490-10	Sequence 10, Appl
30	102	4.2	35	10	US-09-776-491-8	Sequence 8, Appl
31	102	4.2	35	10	US-09-776-491-10	Sequence 10, Appl
32	101.5	4.2	376	9	US-09-738-626-6115	Sequence 115, Ap
33	101	4.2	17	10	US-09-776-490-45	Sequence 45, Appl
34	101	4.2	17	10	US-09-776-490-45	Sequence 45, Appl
35	101	4.2	17	10	US-09-776-491-44	Sequence 44, Appl
36	101	4.2	17	10	US-09-776-491-45	Sequence 45, Appl
37	100	4.1	35	10	US-09-776-490-17	Sequence 17, Appl
38	100	4.1	35	10	US-09-776-491-17	Sequence 17, Appl
39	96	4.0	3069	9	US-09-712-363-246	Sequence 246, App
40	95	3.9	17	10	US-09-776-490-46	Sequence 46, Appl
41	95	3.9	17	10	US-09-776-491-46	Sequence 46, Appl
42	94.5	3.9	282	10	US-09-925-301-1362	Sequence 1362, Ap
43	93	3.9	35	10	US-09-776-490-16	Sequence 16, Appl
44	93	3.9	35	10	US-09-776-491-16	Sequence 16, Appl
45	93	3.9	2201	10	US-09-995-542-9	Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-09-815-242-10253  
Sequence 10253, Application US/09815242  
Patent No. US20020061569A1  
GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsson, Kari L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10253  
LENGTH: 453  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10253  
Query Match 30.5%, Score 734, DB 10, Length 453;  
Best Local Similarity 36.4%, Pred. No. 1, 2e-57;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

```
QY 20 IHGDELFQHELKTIFARNMLFLTHDSLIPAPGDYVYAKMGIDVYASRONDSIRAFIN 79
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 24 IYDDPDYQLELEKIFGSCWLFLEHESQIFPKPGDFNTYMGEDAVYVVRKODGSIKAFIN 83
QY 80 VCRHKGTLVSVENAGKGFVCSYHGWFGSNGELQSVPEFKDLYGESLKNKCLGLKEVA 139
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 84 QCRHRAARVSYADCGNTRAFICPHGMSYGINDELIDVLEPRAYPQGLCSHMGSLNEVP 143
QY 140 RVSEHFHFIYGCDFQEARPLMDYIGDAAMYLEPMF-KHSGLELVGPFGKYVYANKAP 198
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 144 CVESYKCLIGNMDTSPGLRDYLDGIAMYLDMGLDREREGTEIYGVQVWYINCMKFP 203
QY 199 AENFVGAYVHGMTHASSL-----RSGESIFSSLAGNAALP----- 234
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 204 AEOGFASQYIALPASHAAVQYLGAKKDDGSKRLD-----GOTARVWETAKDALOFG 256
QY 235 --DEGAGLQMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKOERLNKEIGDVRARIY 292
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 257 ODGHSGGFETKEKPDAMVWV--DGAVSSYRETYAE-----AEORLGEVRLRL 303
QY 293 RSHLNCIVFPANSMITCSGVFKVNNPIDANTTEWYTAIYVCKMPEDLKRRLADSVOTG 352
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 304 AGHNN--IFPLSLWNLGATLRLVHPRGPDQVEWAFCTDKAASDEVKAAFEENSATRAF 361
QY 353 GPAGFESDNDNMNETSONGKKYQSRDSDLNMGFGEDEVYDAVYGVVYKSAIGETS 412
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 362 GPAGFLEDDSENCELOKLLKHGRANRSKLCLEMGLOQEKRRRDGIGIT--NIIFSETA 420
QY 413 YRGFYRAYQAHVSSSNMAEFEHASSYWHTELTK 445
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 421 ARGMYQRMADLLSSESWQEVLDKTAAYQOEVMK 453
```

## RESULT 2

US-09-815-242-11692  
; Sequence 11692, Application US/09815242  
; Patent No. US20020061569A1

## GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Karl L.  
APPLICANT: Zyskind, Judith W.  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John D.  
APPLICANT: Carr, Grant J.  
APPLICANT: Yamamoto, Robert T.  
APPLICANT: Xu, H. Howard  
TITLE OF INVENTION: Identification of Essential Genes in  
FILE REFERENCE: ELITRA.011A  
CURRENT APPLICATION NUMBER: US/09/815,242  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11692  
LENGTH: 385  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae

US-09-815-242-11692

Query Match 15.7%; Score 377.5; DB 10; Length 385;

Best Local Similarity 31.4%; Pred. No. 5.8e-26;

Matches 113; Conservative 59; Mismatches 145; Indels 43; Gaps 18;

```
QY 34 IFARNMLFLTHDSLIPAPGDYVYAKMGIDVYASRONDSIRAFINVCRRHGTLSVVA 93
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 3 IFEGNWFLAHESQIPQPGDYITTLGRQVYITRDKKNLHLNLINSCARGAMLCRRKT 62
QY 94 GNAGFVCSYHGWFGSNGELQSVPEFK-DLYGESL-NKNCILGLKEVAVSEFHGTYGC 151
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 63 GNKNSFTCPFHGWTFSNNKGLKAKDESTGAYPEPTEHESGSHLOKLRPROSYRGLFQS 122
QY 152 FDOEAPPLMDYIGDAAMYLEPMFKHS--GGLELVGPFGKYVYANKMKAPENFEGDYYHNG 210
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 123 LNAVDPLEAYLGCTCKIILIDVQAPEGLEVLKGSSTIYVEGNWGLKGAEN--GADGYHVS 181
QY 211 ---WTHASSL--RSGESIFSSLAGNAALPPEGAGLQMTSKYG--SGMGVLMDSYSGVHSA 263
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 182 VVHNNVASTMSRRNYEA----EGTHIVDANGMSKSLGGYGFGDNGHMLM-----TR 229
QY 264 DLVPELMAFGAKOERLNKEIGDYRA--RIYRSHLNCIVFPANSMIL--TCSGVFVWNPIDA 321.
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 230 ALNDEVRPV-YAHRERLQAEFGERRADQVNETNLCILYPNVYLMQFSTQIVIRPIAV 288
QY 322 NTFEV--WTYAIY-EKDMPEDLK--RLAD--SVQRTGPGAGFESDNDNME--TASONG 373
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 289 DTEVTYIMCFAPKESQOARALRIROYEDFNVSNGMTP-----DLLEFSAQONG 339
```

## RESULT 3

US-09-738-626-6140  
; Sequence 6140, Application US/09738626  
; Publication No. US20020197605A1

## GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI  
APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, MAKOTO  
APPLICANT: SENO, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 6140  
LENGTH: 490  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum

US-09-738-626-6140

Query Match 15.6%; Score 376; DB 9; Length 490;

Best Local Similarity 26.7%; Pred. No. 1.1e-25;

Matches 121; Conservative 66; Mismatches 208; Indels 58; Gaps 16;

```
QY 20 IHGDELFQHELKTIFARNMLFLTHDSLIPAPGDYVYAKMGIDVYASRONDSIRAFIN 79
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 27 IYDDPDYQLELEKIFGSCWLFLEHESQIFPKPGDFNTYMGEDAVYVVRKODGSIKAFIN 86
QY 80 VCRHKGTLVSVENAGKGFVCSYHGWFGSNGELQSVPEFKD-LYGESLKNK-CLGLKE 137
```

Db 87 ACSHRGMLCRRTDNRTTLTCEPHGWTFSNDGALLKVKDEKEGAYPENFRDGSHLRR 146  
QY 138 VARHESHGFIYCGFDEAPRLMDYLGDAWYLEPMEKHS -GGLELVPPGKVITANKK 196  
Db 147 VPFESTRGTLFSGSLNDVSLSEHIGDTRTVIDMLDVSPEGLLEVLRGSSTVYTDGNMK 206  
QY 197 APAENFEGDAVHGWTHT-----ASSLRGESIFSSLAGNAALPPGAGLOMTSKYSGSM 250  
Db 207 LGQEN-GADQYHVSSTHWNAAATTSRGTGESANETKADAGTWGOGGCFEYPIGHML 265  
QY 251 GVIMDYSGVHSAADVPELMAFGAKOERLKEIGVRRAR-IYRSHLNTCTVPNNMSML-T 308  
Db 266 --LMMWNGNEDRPLF-----ERRDEPKKEFGEKEGFVWGASRNLCLYENVYLMQ 315  
QY 309 CSGVFKWNPIDANTTEWYTAIVEKDMPEDLK----RLADSVORTGCPAGWESDDND 364  
Db 316 FSSQIRIRIRISVDQTEVITYCIAPKGESAEAKANRIROYEDFEFNATG-----MATPD 368  
QY 365 NMEFASONGKKYOSRD---SDLSNLGFGEDVYGDVAVPGVVGK-----SAIGETSY 413  
Db 369 DLEFNSCQCTYQASAFPNWDM--TRGLGHQYOG-----FNEVAKGLGMNEVLTSSGARTD 422  
QY 414 RGFY-----RAYQAHVSSNMWAEFEHAST 438  
Db 423 EGLYPIQHGFWHELMQEAVNKQSIKEKELADDT 455

## RESULT 4

US-09-815-242-5097  
; Sequence 5097, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5097  
; LENGTH: 424  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-5097

Query Match 13.8%; Score:333.5; DB 10; Length 424;  
Best Local Similarity 27.3%; Pred. No. 5 6e-22;  
Matches 114; Conservative 70; Mismatches 178; Indels 55; Gaps 16;  
QY 25 ELFQHELKTIFARNWLFTHDSLIPAPGDYVYAKMGIDEIVSRONDGSIKRAFLNVCRRH 84

Db 24 ELHRRELHEIFEDSMLYAAHLSELREPGDFITRDVGRNLIIOBRADGEPAYVLNACAR 83  
QY 85 GKLTVSEVGNAGKAFVCSYHGMGFSNGELQSVPEPKDLGYSLNKKC---LGLKEVAVY 141  
Db 84 GAKVCAERQNSQRFCCPHGWYDSHSLIGLP-DKAAVQHA--GQCHPELSTQVKRA 140  
QY 142 ESFPHGTYCCFQDEAPRLMDYLGDAWYLEPMEKHS -GGLELVPPGKVITANKAPAE 200  
Db 141 -YRNFLEFHYGARQASLTFTYLGQAKDYIDLICDQSEAELEITIPGFEHSIRANKLAE 199  
QY 201 NPYGDAVHGWTHTASSLRGESIFSSLAGNAALPPGAGLOMTSKYSGSMGLMDYSGV 260  
Db 200 NGV-DAYHLPFAHKRLEYENLTL-----GTPESHKRRGRG-EALGNHALI 244  
QY 261 HS-----ADLVE-LMAFGAKOERLKEIGVRRAR-IYRSHLNTCTVPN--- 303  
Db 245 ISGPPSTGRPIAYWSPLEPALKPSIAAKFERLVERFGQARAEDLAHTNKSIFIPNLVI 304  
QY 304 NSMLTCSGVFKWNPIDANTTEWYTAIVEKDMPEDLKRLADSVORTGCPAGWESDDN 363  
Db 305 NDILGLN--TRSFPTPADEVSYTWGAGFADETREERARINGLISTGPGGFTPDV 362  
QY 364 DNMETASQNGKKYOSRSDLSNLGFGEDVYGDVAVPGVVGKSAIGETSYRGFYRAY 420  
Db 363 ELIESQ---RAYAH-----AALGYSDFSRG---MGPATRRHVDENRGFWREM 406

## RESULT 5

US-09-776-490-19  
; Sequence 19, Application US/09776490  
; Patent No. US20010012886A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; APPLICANT: John, Gurmukh S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009  
; CITY: Charlotte  
; STATE: No. US20010012886A1th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/776,490  
; FILING DATE: 02-Feb-2001  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/810,009  
; FILING DATE: 04-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spullin, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: 5718-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; TELE: 575102  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-490-19

Query Match 7.7%: Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%: Pred. No. 2.7e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGSNGE 113  
Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGSNGK 35

RESULT 6  
US-09-776-491-19

Sequence 19, Application US/09776491  
Patent No. US20010013135A1

GENERAL INFORMATION:

APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.

Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009

CITY: Charlotte

STATE: No. US20010013135A1th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/776,491

FILING DATE: 02-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/810,009

FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sprull, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: 5718-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175

TELEX: 575102

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 35 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-491-19

Query Match 7.7%: Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%: Pred. No. 2.7e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGSNGE 113  
Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGSNGK 35

RESULT 7  
US-09-776-490-20

Sequence 20, Application US/09776490  
Patent No. US20010012886A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.

Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009

CITY: Charlotte

STATE: No. US20010012886A1th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/776,490

FILING DATE: 02-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/810,009

FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sprull, W. Murray

REGISTRATION NUMBER: 32,943

REFERENCE/DOCKET NUMBER: 5718-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175

TELEX: 575102

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 35 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-490-20

Query Match 7.2%: Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%: Pred. No. 2.5e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGSNGE 113  
Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGSNGK 35

RESULT 8  
US-09-776-491-20

Sequence 20, Application US/09776491  
Patent No. US20010013135A1

GENERAL INFORMATION:

APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.

Gray, John

Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65

CORRESPONDENCE ADDRESS:

ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009

CITY: Charlotte

STATE: No. US20010013135A1th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/776,491  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-491-20

Query Match 7.2%; Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%; Pred. No. 2.5e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWGFGSNGK 35

99 NVCRHRGKTLVSEAGNAKGFVCSYHGWGFGSNGE 113

RESULT 9  
US-09-776-490-21  
Sequence 21, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-490-21

Query Match 7.0%; Score 168; DB 10; Length 35;  
Best Local Similarity 80.0%; Pred. No. 8.7e-09;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Db 1 NVCRHRGKTLVDAEAGNAKGPVCGYHGWGFGSNGK 35

99 NVCRHRGKTLVSEAGNAKGFVCSYHGWGFGSNGE 113

RESULT 10  
US-09-776-491-21  
Sequence 21, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-491-21

Query Match 7.0%; Score 168; DB 10; Length 35;  
Best Local Similarity 80.0%; Pred. No. 8.7e-09;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCRRHKGTLVSVAGNAKGVCVSYHGMGFGSNGE 113  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 NVCRRHKGTLVDAAGNAKGVCVSYHGMGFGSNGK 35

RESULT 11  
US-08-976-063C-4  
; Sequence 4, Application US/08976063C  
; Publication No. US20020182697A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Steinduchel, Horst Priefert, Jürgen Rabenhorst  
; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF  
; TITLE OF INVENTION: CONIFERYL ALCOHOL, CONIFERYLALDEHYDE, FERULIC ACID, VANILLIN A  
; TITLE OF INVENTION: ACID AND THEIR USE  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-5144  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage  
; OPERATING SYSTEM: DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/976,063C  
; FILING DATE: 21-NOV-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 196 49 655.1 (Germany)  
; FILING DATE: 29-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141  
; REFERENCE/DOCKET NUMBER: Bayer 9998-CAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 332-1700  
; TELEFAX: (914) 332-1844  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 354 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-976-063C-4

Query Match 4.9%; Score 118.5; DB 1; Length 354;  
Best Local Similarity .23.2%; Pred. No. 0.0063;  
Matches 56; Conservative 40; Mismatches 78; Indels 67; Gaps 14;  
QY 34 IFARN-WFLTHDSLIPADGVYTAAMGIDEIVTSQNDGSIKRAFLNVCRRHKGTLVSV 91  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 MFKKNMAYVACTDEIDAKP---LGRQICNKEKIVFYGRGREGRAVAVEDFCRHGRAPL--- 54  
QY 92 EAGNARGFV-----CSYHGMGFGSNGELOSVPREKDLVGSLLKKKLG---LKEVARVE 142  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 55 ----SLGFVNDGKILGCGHLEMGCEGKTLAMP-----GQRYOGFPCIKSYA-VE 99  
QY 143 SFHGFTY---GCFDQEARPPMDVLDGDAWYLEPFRKHSGLLELVGPPGVYIKANMKAPA 199  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 100 ERGCFITWVGDRDLRALIHNL---EMADNPEMAIGGGL-----YIIADYRLMT 148  
QY 200 ENFVDAYHVGWTHASSLR-----SGESIFSS-LAGNAALRP-----EGAG 239  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 149 DNLML-DLHETTYHASSIGCKEIDEAPVSTRVEDGVITTSYMDNVNAPRPMRAALRGNG 207  
QY 240 L 240  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 208 L 208

RESULT 12  
US-09-776-490-12  
; Sequence 12, Application US/09776490  
; Patent No. US20010012886A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; Jhal, Gurmukh S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009  
; CITY: Charlotte  
; STATE: No. US20010012886A1th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/776,490  
; FILING DATE: 02-Feb-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/810,009  
; FILING DATE: 04-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spruill, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: 5718-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; TELEX: 575102  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-490-12  
Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0013;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;  
QY 79 NVCRRHKGTLVSVAGNAKGVCVSYHGMGFGSNG 112  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 NVCRRHGRICRADAGNAKAPCTCSYHGMAYDTAG 34  
RESULT 13  
US-09-776-491-12  
; Sequence 12, Application US/09776491  
; Patent No. US20010013135A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; Jhal, Gurmukh S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009

/? CITY: Charlotte  
/? STATE: No. US20010013135A1th Carolina  
/? COUNTRY: USA  
/? ZIP: 28234  
/?  
/? COMPUTER READABLE FORM:  
/? MEDIUM TYPE: floppy disk  
/? COMPUTER: IBM PC compatible  
/? OPERATING SYSTEM: PC-DOS/MS-DOS  
/? SOFTWARE: Patentin Release #1.0, Version #1.30  
/? CURRENT APPLICATION DATA:  
/? APPLICATION NUMBER: US/09/776,491  
/? FILING DATE: 02-Feb-2001  
/? CLASSIFICATION: <Unknown>  
/? PRIOR APPLICATION DATA:  
/? APPLICATION NUMBER: US 08/810,009  
/? FILING DATE: 04-MAR-1997  
/? ATTORNEY/AGENT INFORMATION:  
/? NAME: Spruill, W. Murray  
/? REGISTRATION NUMBER: 32,943  
/? REFERENCE/DOCKET NUMBER: 5718-4  
/? TELECOMMUNICATION INFORMATION:  
/? TELEPHONE: 919-881-3140  
/? TELEFAX: 919-881-3175  
/? TELEX: 575102  
/? INFORMATION FOR SEQ ID NO: 12:  
/? SEQUENCE CHARACTERISTICS:  
/? LENGTH: 35 amino acids  
/? TYPE: amino acid  
/? STRANDEDNESS: <unknown>  
/? TOPOLOGY: linear  
/? MOLECULE TYPE: protein  
/? SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-491-12  
  
Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0013;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;  
  
QY 79 NVCRHKGKTLVSEAGNAGFCVSYHGFGSNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NCCRHRGMRICRADGNAGNAKAPTCSYHGWAYVDG 34  
  
RESULT 14  
US-09-776-490-14  
/? Sequence 14, Application US/09776490  
/? Patent No. US20010012886A1  
/? GENERAL INFORMATION:  
/? APPLICANT: Briggs, Steven P.  
/? John, Gurmukh S.  
/? Gray, John  
/? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
/? CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
/? NUMBER OF SEQUENCES: 65  
/? CORRESPONDENCE ADDRESS:  
/? ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
/? STREET: P.O. Drawer 34009  
/? CITY: Charlotte  
/? STATE: No. US20010012886A1th Carolina  
/? COUNTRY: USA  
/? ZIP: 28234  
/? COMPUTER READABLE FORM:  
/? MEDIUM TYPE: floppy disk  
/? COMPUTER: IBM PC compatible  
/? OPERATING SYSTEM: PC-DOS/MS-DOS  
/? SOFTWARE: Patentin Release #1.0, Version #1.30  
/? CURRENT APPLICATION DATA:  
/? APPLICATION NUMBER: US/09/776,490  
/? FILING DATE: 02-Feb-2001  
/? CLASSIFICATION: <Unknown>  
/? PRIOR APPLICATION DATA:  
/? APPLICATION NUMBER: US 08/810,009  
/? FILING DATE: 04-MAR-1997

/? ATTORNEY/AGENT INFORMATION:  
/? NAME: Spruill, W. Murray  
/? REGISTRATION NUMBER: 32,943  
/? REFERENCE/DOCKET NUMBER: 5718-4  
/? TELECOMMUNICATION INFORMATION:  
/? TELEPHONE: 919-881-3140  
/? TELEFAX: 919-881-3175  
/? TELEX: 575102  
/? INFORMATION FOR SEQ ID NO: 14:  
/? SEQUENCE CHARACTERISTICS:  
/? LENGTH: 35 amino acids  
/? TYPE: amino acid  
/? STRANDEDNESS: <unknown>  
/? TOPOLOGY: linear  
/? MOLECULE TYPE: protein  
/? SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-776-490-14  
  
Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0016;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
  
QY 79 NVCRHKGKTLVSEAGNAGFCVSYHGFGSNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NCCRHRGMRICRADGNAGNAKAPTCSYHGWAYVDG 34  
  
RESULT 15  
US-09-776-491-14  
/? Sequence 14, Application US/09776491  
/? Patent No. US20010013135A1  
/? GENERAL INFORMATION:  
/? APPLICANT: Briggs, Steven P.  
/? John, Gurmukh S.  
/? Gray, John  
/? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
/? CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
/? NUMBER OF SEQUENCES: 65  
/? CORRESPONDENCE ADDRESS:  
/? ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
/? STREET: P.O. Drawer 34009  
/? CITY: Charlotte  
/? STATE: No. US20010013135A1th Carolina  
/? COUNTRY: USA  
/? ZIP: 28234  
/? COMPUTER READABLE FORM:  
/? MEDIUM TYPE: floppy disk  
/? COMPUTER: IBM PC compatible  
/? OPERATING SYSTEM: PC-DOS/MS-DOS  
/? SOFTWARE: Patentin Release #1.0, Version #1.30  
/? CURRENT APPLICATION DATA:  
/? APPLICATION NUMBER: US/09/776,491  
/? FILING DATE: 02-Feb-2001  
/? CLASSIFICATION: <Unknown>  
/? PRIOR APPLICATION DATA:  
/? APPLICATION NUMBER: US 08/810,009  
/? FILING DATE: 04-MAR-1997  
/? ATTORNEY/AGENT INFORMATION:  
/? NAME: Spruill, W. Murray  
/? REGISTRATION NUMBER: 32,943  
/? REFERENCE/DOCKET NUMBER: 5718-4  
/? TELECOMMUNICATION INFORMATION:  
/? TELEPHONE: 919-881-3140  
/? TELEFAX: 919-881-3175  
/? TELEX: 575102  
/? INFORMATION FOR SEQ ID NO: 14:  
/? SEQUENCE CHARACTERISTICS:  
/? LENGTH: 35 amino acids  
/? TYPE: amino acid  
/? STRANDEDNESS: <unknown>  
/? TOPOLOGY: linear  
/? MOLECULE TYPE: protein  
/? SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-776-491-14

Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0016;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY 79 NVCRRHCKTLVSEAGNAKGFVCSYHGMGFGSNG 112  
| | | | | : | | | | | : | |  
Db 1 NQCRHRGMRICRADCGNAKSPFCSYHGNAYDSAG 34

Search completed: January 27, 2003, 09:20:55  
Job time : 8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:56:07 ; Search time 10.1667 Seconds  
(without alignments)  
1299.432 Million cell updates/sec

Title: US-09-843-250-32

Perfect score: 2410  
Sequence: 1 MNYNKKILVSEGLSOKHLI.....AEFEHASSWHTLTKTTDR 449

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	224	9.3	446	4	US-09-004-393B-4
2	223	9.3	439	4	US-09-004-393B-2
3	185	7.7	35	4	US-08-810-009-19
4	174	7.2	35	4	US-08-810-009-20
5	168	7.0	35	4	US-08-810-009-21
6	113.5	4.7	379	3	US-09-028-934-36
7	110	4.6	35	4	US-08-810-009-12
8	109	4.5	35	4	US-08-810-009-14
9	108	4.5	35	4	US-08-810-009-13
10	108	4.5	35	4	US-08-810-009-15
11	108	4.5	622	4	US-09-311-626B-4
12	106	4.4	35	4	US-08-810-009-9
13	105.5	4.4	432	4	US-08-809-326A-16
14	105.5	4.4	649	4	US-08-809-326A-15
15	104	4.3	35	4	US-08-810-009-18
16	103	4.3	35	4	US-08-810-009-11
17	102	4.2	35	4	US-08-810-009-8
18	102	4.2	35	4	US-08-810-009-10
19	101	4.2	17	4	US-08-810-009-44
20	101	4.2	17	4	US-08-810-009-45
21	100	4.1	35	4	US-08-810-009-17
22	99.5	4.1	256	4	US-09-325-932A-57
23	99	4.1	560	3	US-08-814-052-6
24	99	4.1	560	3	US-08-812-829-6
25	95	3.9	17	4	US-08-810-009-46
26	95	3.9	563	4	US-09-134-001C-4800
27	94.5	3.9	350	2	US-08-828-922-1

28	94	3.9	5588	4	US-09-036-987A-6	Sequence 6, Appl1
29	94	3.9	5588	4	US-09-370-700-6	Sequence 6, Appl1
30	93	3.9	35	4	US-08-810-009-16	Sequence 16, Appl1
31	93	3.9	421	4	US-09-239-303-2	Sequence 2, Appl1
32	92	3.8	1087	2	US-08-570-311-8	Sequence 8, Appl1
33	92	3.8	1087	2	US-08-353-485-8	Sequence 8, Appl1
34	92	3.8	1358	2	US-08-570-311-27	Sequence 27, Appl1
35	91.5	3.8	525	1	US-08-348-891A-2	Sequence 2, Appl1
36	91.5	3.8	525	2	US-08-905-817-2	Sequence 2, Appl1
37	91.5	3.8	750	3	US-08-814-052-2	Sequence 2, Appl1
38	91.5	3.8	750	3	US-08-812-829-2	Sequence 2, Appl1
39	91	3.8	471	4	US-08-914-375C-70	Sequence 70, Appl1
40	90.5	3.8	345	4	US-09-362-123A-2	Sequence 2, Appl1
41	90.5	3.8	346	4	US-09-149-476-493	Sequence 493, App
42	90.5	3.8	774	4	US-09-276-400-7	Sequence 7, Appl1
43	90.5	3.8	774	4	US-09-448-076-7	Sequence 7, Appl1
44	90.5	3.8	774	4	US-09-702-572-7	Sequence 7, Appl1
45	90.5	3.8	997	1	US-08-232-540-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-09-004-393B-4
; Sequence 4, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
; APPLICANT: Rathnasabapathi, Bala
; APPLICANT: Burnet, Michael
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; FILE REFERENCE: UF-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; CURRENT FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-004-393B-4

Query Match          9.3%; Score 224; DB 4; Length 446;
Best Local Similarity 27.4%; Pred. 13;
Matches 58; Conservative 37; Mismatches 97; Indels 20; Gaps 6;

QY 11 ESGLSKHLIHGDELFOHELTIFARNMLFLTHDSLIPAPGDYTAAGKIDIEVIYSRON 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 99 EALTPPSTWYTEPARYSHLEIFYKQWVAGYSPOVEKNOYFGSLGNEYLVSRDG 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 71 DGSIRFLVWCRGRKTVLSVEAGNAKGFVCSYHGFGSNGELSVPREKLDYGSINK 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 159 QGELIAHFVHCNHRASILACSGSKSCFCVCRHGVNYLDLSLAKA--SKATEPQNDLP 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 KCLIGLEKAVESFHGFIYGCEDQEAAPLMD----YLGA-----AWYLEPMFKHSGCLE 181
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 KETGLAPL-KVAEWGPFILISDRSIDANADVTEWIKSADYKANAHPDNLKFTHRSE 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 LVGPPEKVIYKANWKAAPENFPGDAYHVQWTH 213
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 F-----PMECNMKVFCDNYLDSSYHVPYAH 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-09-004-393B-2
; Sequence 2, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
```

```

; APPLICANT: Rathinasabapathi, Bala
; APPLICANT: Burnet, Michael
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; TITLE OF INVENTION: Plants Transformed Therewith
; FILE REFERENCE: US-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; CURRENT FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 439
; TYPE: PRN
; ORGANISM: Splachia oleracea
; US-09-004-393B-2

Query Match
Best Local Similarity 9.3%; Score 223; DB 4; Length 439;
Matches 58; Conservative 38; Mismatches 83; Indels 14; Gaps 7;

Qy 27 FOHEKTFEANNLELTHDSILIPAGDYVTAKMGIDEVIVSQNGSIRAFINCRHRCK 86
Db 108 YSHELRLFYGMQVAGISDQIKENQFTGSLGNVEYLVSHDGGKVAHFNVCTHRA- 166
Qy 87 TLVVEAGNAGKFCVSYHGMFGSNGELQSVPEEKDLVGESLNKCKLGLKEVARVESFHG 146
Db 167 SILAGSGKSCFCVPHYGMVYGMGSLAKASKAP--EQNDPHELGLVPL-KVAVWCP 223
Qy 147 FYVCFDQAEPLMD---YLGDAAWLEPMFKHS--GGLVLPGGKVIKANKKAPAE 200
Db 224 FVLISLRLEEGDVGTEMLGTSA---EDVKAHAFDPSLQFI-HRSEPPMESNMKIFSD 279
Qy 201 NFVGDAYHVGWTH 213
Db 280 NYLDSYHVPYAH 292

RESULT 3
US-08-810-009-19
; Sequence 19, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; City: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. MurRAY
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
```

```

; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-19

Query Match
Best Local Similarity 7.7%; Score 185; DB 4; Length 35;
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 79 NVCRRGKTLVSEAGNAGKFCVSYHGMFGSNGE 113
Db 1 NVCRRGKTLVSEAGNAGKFCVSYHGMFGSNGK 35

RESULT 4
US-08-810-009-20
; Sequence 20, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; City: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. MurRAY
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-20

Query Match
Best Local Similarity 7.2%; Score 174; DB 4; Length 35;
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 79 NVCRRGKTLVSEAGNAGKFCVSYHGMFGSNGE 113
Db 1 NVCRRGKTLVSEAGNAGKFCVSYHGMFGSNGK 35

RESULT 5
US-08-810-009-21
```

```
; Sequence 21, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spurrill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-21

Query Match      7.0%; Score 168; DB 4; Length 35;
Best Local Similarity 80.0%; Pred. No. 1.2e-09;
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCHRGRKTIYSVAGNAKGFVCSYHGMGSGNCE 113
Db 1 NVCHRGRKTIYDAAGNAKGPVCGYHGMGYSNGK 35

RESULT 6
US-09-028-934-36
; Sequence 36, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kliner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyroclitilin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-028-934-36

Query Match      4.7%; Score 113.5; DB 3; Length 379;
Best Local Similarity 20.3%; Pred. No. 0.0098;
Matches 95; Conservative 54; Mismatches 149; Indels 171; Gaps 26;
```

```
QY 20 IHGDEELFQ----HEKITYFARNW-LFTHDSLIPAPDDYVTAKKGIDEIVYSRQDGS 73
Db 5 IHQEPRIQASGVNDLTQTASWYVAMRSDALRGKP--VAIKLFGQPLVAMRDGGGR 61

QY 74 IRAFLWVCRGRKTIYSVEAGNAKGFV-----CSYHGMGFGSNGELQSVPEPKLDYGS 127
Db 62 PYVMERYCSHLGASL-----AKGKVVEGCTQCPTPHNMRYDSTGACSHVP-----GHS 108

QY 128 LNKKGLCKEVARVESF-----HGFIYGFCEQDEAP-----PLMDYLGDAAY 169
Db 109 -----TEVRLPELEIPTAROSVYPMVERGFGVWVWGTARLPPLPEMEAEBSSEH 160

QY 170 LEPMEFHSGLGLVPPGKVYIKANKKAPADENFVGDAYVGMTHASSLSRGE-SIF---S 225
Db 161 QSLRFAYETTTSVL---RRI-----ENFY-DAQHAAPVHQLPISAFELKLFDESS 206

QY 226 SLAGNALPPE---GAGLQW-TSKYSGMGVL-----WDGSGVHSADLV 266
Db 207 PPPGQALARDGAMFGAGIDFHVDRIFGPLGVISRTLGISMSBMOIHFGYFG----- 259

QY 267 PELMAFGAKOERLNKEIDVARIYRSHLNCTVPPNNSMLTSCGVFKVWNPIDANTTEV 326
Db 260 -----GCIMTVSLD---GDVKYRL---LQCV-----TPVDKEETVW 289

QY 327 WTYAIYEKMPEDLKRRLADSV---QRTGCPAGF---WESDNDNMEFASONGKKYOSR 379
Db 290 HMLAIKKG--DGVVRSANFLIYGLQTTWAAGYDVAIINSMKAD-----GGAGAFSKY 340

QY 380 DSDLLSNLGFEGDYGDVAYPGVYKSAIGENSYSGFYRAYAYSSN 428
Db 341 DDLIL-----KYRAFYRWKVALEN 362

RESULT 7
US-08-810-009-12
; Sequence 12, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
```

```

1 CLASSIFICATION: 800
2 ATTORNEY/AGENT INFORMATION:
3 NAME: Sproull, W. Murray
4 REGISTRATION NUMBER: 32,943
5 REFERENCE/DOCKET NUMBER: 5718-4
6 TELECOMMUNICATION INFORMATION:
7 TELEPHONE: 919-881-3140
8 TELEFAX: 919-881-3175
9 TELEX: 575102
10 INFORMATION FOR SEQ. ID NO: 14:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 35 amino acids
13 TYPE: amino acid
14 STRANDEDNESS:
15 TOPOLOGY: linear
16 MOLECULE TYPE: protein
17 US-08-810-009-14
18
19 Query Match 4.5%; Score 109; DB 4; Length 35;
20 Best Local Similarity 52.9%; Pred. No. 0.00067;
21 Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
22
23 Qy 79 NVCHRGKTLVSEAGNAKPGVCSYHGNGFGSG 112
24 1 11111 : : 1111 11111 : 1
25 Db 1 NQCHRGMRICRADGNAKSPCTSYHGNAVDAG 34
26
27 RESULT 9
28 US-08-810-009-13
29 Sequence 13, Application US/08810009
30 Patent No. 6211437
31 GENERAL INFORMATION:
32 APPLICANT: Briggs, Steven P.
33 APPLICANT: Johal, Gurmukh S.
34 APPLICANT: Gray, John
35 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
36 NUMBER OF SEQUENCES: 65
37 TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
38 CORRESPONDENCE ADDRESS:
39 ADDRESSEE: BELL, SELTZER, PARK & GIBSON
40 STREET: P.O. Drawer 34009
41 City: Charlotte
42 STATE: No. 6211437th Carolina
43 COUNTRY: USA
44 ZIP: 28234
45 COMPUTER READABLE FORM:
46 MEDIUM TYPE: Floppy disk
47 COMPUTER: IBM PC compatible
48 OPERATING SYSTEM: PC-DOS/MS-DOS
49 SOFTWARE: PatentIn Release #1.0, Version #1.30
50 CURRENT APPLICATION DATA:
51 APPLICATION NUMBER: US/08/810,009
52 FILING DATE: 04-MAR-1997
53 CLASSIFICATION: 800
54 ATTORNEY/AGENT INFORMATION:
55 NAME: Sproull, W. Murray
56 REGISTRATION NUMBER: 32,943
57 REFERENCE/DOCKET NUMBER: 5718-4
58 TELECOMMUNICATION INFORMATION:
59 TELEPHONE: 919-881-3140
60 TELEFAX: 919-881-3175
61 TELEX: 575102
62 INFORMATION FOR SEQ. ID NO: 13:
63 SEQUENCE CHARACTERISTICS:
64 LENGTH: 35 amino acids
65 TYPE: amino acid
66 STRANDEDNESS:
67 TOPOLOGY: linear
68 MOLECULE TYPE: protein
69 US-08-810-009-13
70
71 Query Match 4.5%; Score 108; DB 4; Length 35;
72 Best Local Similarity 52.9%; Pred. No. 0.00084;

```



	Matches	18;	Conservative	4;	Mismatches	12;	Indels	0;	Gaps	0
Qy	79	NVCRHRGKTIYSVEAGNAKGIFYCSYHGMFGSANG	112							
		:	:	:	:	:	:	:	:	:
Db	1	NQCRHRGMRICSDAGNAKAPTCYSYHGNAIYDAG	34							

US-08-810-009-15  
RESULT 10  
Sequence 15, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Jonal, Gurmukh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
City: Charlotte  
STATE: NO. 6211437b Carolina  
COUNTRY: USA  
ZIP: 28234

Query Match	4.5%	Score 108	DB 4	Length 35
Best Local Similarity	50.0%	Pred. No. 0.00084		
Matches 17; Conservative		4; Mismatches 13; Indels	0; Gaps	0;
Qy	79	NVCRHGRKTLVSEAGNAKPFVCSITHGFGFSNG	112	
				.
Ob	1	NCRHGRKICRADGGMASPTCSYHGAAVYTG	34	

RESULT 11  
 US-09-311-6266-4  
 ; Sequence 4, Application US/09311626B  
 ; Patent No. 6399347  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jorgensen, Per Lina  
 ; APPLICANT: Schmoor, Kirk  
 ; APPLICANT: Andersen, Lene No. 6399347b0e  
 ; APPLICANT: Schulteib, Martin  
 ; APPLICANT: Outtrup, Helge  
 ; TITLE OF INVENTION: No. 6399347e1 Rhamnogalacturonan Hydrolases  
 ; FILE REFERENCE: 5572.204-US  
 ; CURRENT APPLICATION NUMBER: US/09/311,626B

```

:
: CURRENT FILING DATE: 1999-05-13
: PRIOR APPLICATION NUMBER: 0608/98
: PRIOR FILING DATE: 1998-05-01
: PRIOR APPLICATION NUMBER: 60/084,358
: PRIOR FILING DATE: 1998-05-05
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 622
: TYPE: prt
: ORGANISM: Bacillus licheniformis
: US-09-311-626B-4

```

[illegible]

RESULT 12  
 US-08-810-009-9  
 Sequence 9, Application US/08810009  
 Patent No. 6211437  
 GENERAL INFORMATION:  
 APPLICANT: Briggs, Steven P.  
 APPLICANT: Johal, Gurmukh S.  
 APPLICANT: Gray, John  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
 TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
 STREET: P.O. Drawer 34009  
 CITY: Charlotte  
 STATE: No. 6211437/1b Carolina  
 COUNTRY: USA  
 ZIP: 28234  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/810,009  
 FILING DATE: 04-MAR-1997  
 CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-9

Query Match  
Best Local Similarity 51.4%; Score 106; DB 4; Length 35;  
Matches 18; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

Db 79 NCRHRGKTLVSEAGNAKGFCSYHGMGFGSNGE 113  
1 NCRHRGMRICRSDAGNAKAPTCSYHGMAYDIACK 35

RESULT 13  
US-08-809-326A-16  
Sequence 16, Application US/08809326A  
Patent No. 6165478  
GENERAL INFORMATION:  
APPLICANT: Izutsu, Hiroshi  
APPLICANT: Obata, Kazuhiko  
APPLICANT: Matsumoto, Akira  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMANTS  
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAs CODING  
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMANTS  
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF  
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR  
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
PNEUMONIAE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106009/95  
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106010/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106011/95  
FILING DATE: 28-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Charles E.  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 7426-043-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-809-326A-16

Query Match  
Best Local Similarity 21.4%; Score 105.5; DB 4; Length 432;  
Matches 53; Conservative 38; Mismatches 90; Indels 67; Gaps 11;

Db 195 WKAPAEENVGDYHVGWTHASSL-----RSGESIFSSLAGNALPPEGAGLQWTSKY 246  
22 WNLPAD-----LAMFKRNTLKPVMGRHTWESIGRLPR-----KNIIISQP 66  
QY 247 GSGMGLMDGYSVSHADL-----VPELMAFGAKOERLNEIGDVARIRYSHLNC 299  
Db 67 GTDDRTW-----KVSDEALIAAGDVEIIMVIGG---RVYEQPLPKQOKLYLTHIDAE 118  
QY 300 V-----PENNMLTCSGVFRVWNPIDANTTEVTVYAVEDEMDKRLRLASVORTGP 354  
Db 119 VEGDTHPDVPEDDWESFEFHDAQNHSYEFELT-----RLNLSISSSGP 170  
QY 355 AGEWESDDNNME---TASQNGKKYQSRSD-----LLSNLFGEDYGDYGVYGVG 404  
Db 171 -----DNQNIINSQVLTSPQGVPPQDKLSGNETKQIQOTRGKNTENESDATTAGAG 224  
QY 405 KSAIGETS 412  
Db 225 KDKTSST 232

RESULT 14  
US-08-809-326A-15  
Sequence 15, Application US/08809326A  
Patent No. 6165478  
GENERAL INFORMATION:  
APPLICANT: Izutsu, Hiroshi  
APPLICANT: Obata, Kazuhiko  
APPLICANT: Matsumoto, Akira  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMANTS  
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAs CODING  
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMANTS  
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF  
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
PNEUMONIAE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711

```

? APPLICANT: Biggs, Steven P.
? APPLICANT: Johal, Gurmukh S.
? APPLICANT: Gray, John
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
? TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
? NUMBER OF SEQUENCES: 65
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BELL, SELTZER, PARK & GIBSON
? STREET: P.O. Drawer 34009
? CITY: Charlotte
? STATE: No. 6211437th Carolina
? COUNTRY: USA
? ZIP: 28234
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/810,009
? FILING DATE: 04-MAR-1997
? CLASSIFICATION: 800
? ATTORNEY/AGENT INFORMATION:
? NAME: Spurrill, W. Murray
? REGISTRATION NUMBER: 32,943
? REFERENCE/DOCKET NUMBER: 5718-4
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-881-3140
? TELEFAX: 919-881-3175
? TELEX: 575102
? INFORMATION FOR SEQ ID NO: 18:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 35 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-810-009-18

Query Match          4.3%; Score 104; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.0021;
Matches 18; Conservative 2; Mismatches 14; Indels 0; Gaps 0

QY      79 NVCRRHGKTLVSEAGNAKGFCVSYHGMGFGSGNG 112
        I IIII I II I III I : I I
Db       1 NSCRHRGALLCPFSKGNOKFHVCRYHGWMSYSSG 34

Search completed: January 27, 2003, 09:04:53
Job time : 11.1667secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:54:32 ; Search time 25.8333 Seconds  
(without alignments)  
3581.232 Million cell updates/sec

Title: US-09-843-250-32  
Perfect score: 2410  
Sequence: 1 MNYNKKILVSESGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2329	96.6	449	2	Q938R9
2	2289	95.0	449	2	Q52142
3	2205	91.5	449	2	Q92173
4	2190	90.9	447	2	Q915D2
5	2186	90.7	447	2	Q52382
6	2068	85.8	447	2	Q8VUD4
7	2048	85.0	447	2	P95564
8	2019	83.8	447	2	Q8RTI4
9	1966.5	81.6	451	2	Q45695
10	1490	61.8	277	2	Q9ETR2
11	1488	61.7	277	2	Q9F5S4
12	1486	61.7	277	2	Q9F5T3
13	1484	61.6	277	2	Q9F5S8
14	1483	61.5	277	2	Q9F5S9
15	1451	60.2	277	2	Q9F5S3
16	1405	58.3	277	2	Q9EUC7

17	1403	58.2	277	2	Q9F5T8	Q9F5T8 pseudomonas
18	1402	58.2	277	2	Q9F5T5	Q9F5T5 pseudomonas
19	1399	58.0	277	2	Q9ETV3	Q9ETV3 pseudomonas
20	1399	58.0	277	2	Q9F5S5	Q9F5S5 pseudomonas
21	1396	57.9	277	2	Q9F5T6	Q9F5T6 pseudomonas
22	1396	57.9	277	2	Q9F5S7	Q9F5S7 pseudomonas
23	1393	57.8	277	2	Q9F5T4	Q9F5T4 pseudomonas
24	1393	57.8	277	2	Q9F5T2	Q9F5T2 pseudomonas
25	1393	57.8	277	2	Q9F5T1	Q9F5T1 pseudomonas
26	1393	57.8	277	2	Q9F5T0	Q9F5T0 pseudomonas
27	1393	57.8	277	2	Q9F5S6	Q9F5S6 pseudomonas
28	1390	57.7	277	2	Q9F5T7	Q9F5T7 pseudomonas
29	1384	57.4	277	2	Q9F5S2	Q9F5S2 pseudomonas
30	1352.5	56.1	256	2	Q9ZHH3	Q9ZHH3 burkholderi
31	1139	47.3	226	2	Q8VUM6	Q8VUM6 ralstonia s
32	1134	47.1	214	2	Q8VR23	Q8VR23 pseudomonas
33	1134	47.1	214	2	Q8VL21	Q8VL21 pseudomonas
34	1127	46.8	214	2	Q8VR24	Q8VR24 pseudomonas
35	1123	46.6	214	2	Q8VR22	Q8VR22 pseudomonas
36	1110.5	46.1	455	2	Q93NA8	Q93NA8 burkholderi
37	1067	44.3	214	2	Q8VR25	Q8VR25 burkholderi
38	1042.5	43.3	437	2	Q9WKG8	Q9WKG8 alcaligenes
39	1040	43.2	208	2	Q9F6B6	Q9F6B6 marinobacte
40	1039	43.1	206	2	Q8VUM5	Q8VUM5 naphthalene
41	1038	43.1	214	2	Q8VR21	Q8VR21 burkholderi
42	1038	43.1	303	2	Q8VUM2	Q8VUM2 ralstonia s
43	1016	42.2	297	2	Q93MA0	Q93MA0 uncultured
44	1009	41.9	297	2	Q93M39	Q93M39 uncultured
45	1002	41.6	297	2	Q93M41	Q93M41 uncultured

## ALIGNMENTS

RESULT 1	ID	Q938R9	PRELIMINARY:	PRT:	449 AA.
AC	Q938R9:	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Naphthalene dioxygenase.				
GN	NAHAC.				
OS	Pseudomonas fluorescens.				
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;				
CC	Pseudomonas.				
OX	NCBI_TaxID=294;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Min K.-H., Ji S.-H.;				
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AY048759; AAL07262.1; -				
DR	InterPro; IPR001281; Rlske.				
DR	InterPro; IPR001663; Ring_hydroxyl_A.				
DR	Pfam; PF00355; Rlske; 1.				
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.				
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.				
KW	Dioxygenase.				
SO	SEQUENCE 449 AA; 49608 MW; 5EEDBE0282FE812 CRC64;				
Query Match	96.6%; Score 2329; DB 2; Length 449;				
Best Local Similarity	96.0%; Pred. No. 4e-171;				
Matches	431; Conservative 12; Mismatches 6; Indels 0; Gaps 0;				
Qy	1	MNYNKKILVSESGLSQKHLIHGDEELFOHELKTIFARWMLFTHDSLIAPAGDYVTAKG 60			
Db	1	MNYNKKILVSESGLSQKHLIHGDEELFOHELKTIFARWMLFTHDSLIAPAGDYVTAKG 60			
Qy	61	IDENVSRONGSTRAFLNVCGRHKGTLVSVEAGNAKGFVCSYHNGMGSGNCELOSYPFE 120			
Db	61	IDENVSRONGSTRAFLNVCGRHKGTLVSVEAGNAKGFVCSYHNGMGSGNCELOSYPFE 120			
Qy	121	KDLVGSINRKCILGEVARVESFNGFIYCGFDGRAPLMDYLDGDAWYLIPMFKHSGL 180			

```

      121 KELYGESLNKCLGLKEVARSFHGFYCGFDEAPSLMDYLGDAAYLPIFKHSGL 180
      181 ELVGPPKVVYIKAMWKAPEAFNFVGDAHYVGMTTHASSLRSGESIFSSLAGNALPPREGAGL 240
      181 ELVGPPKVVYIKAMWKAPEAFNFVGDAHYVGMTTHASSLRSGESIFSSLAGNALPPREGAGL 240
      241 QMTSKYSGMGVLWDGYSVHSADLVPELMAFGAKOERLNKEIGDVARIRYRSHLNCV 300
      241 QMTSKYSGMGVLWDGYSVHSADLVPELMAFGAKOERLNKEIGDVARIRYRSHLNCV 300
      241 QMTSKYSGMGVLWDGYSVHSADLVPELMAFGAKOERLNKEIGDVARIRYRSHLNCV 300
      301 FPNNSVLTGSGVFKVWNPIDANTTEVNTYALVEKDMPEDLKRLADAVQRTFGPAGFWS 360
      301 FPNNSVLTGSGVFKVWNPIDANTTEVNTYALVEKDMPEDLKRLADAVQRTFGPAGFWS 360
      361 DDNDNMTASONGKRYKSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
      361 DDNDNMTASONGKRYKSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
      421 QAHVSSSNMAEFEDASSTWHTELTKTTDR 449
      421 QAHVSSSNMAEFEDASSTWHTELTKTTDR 449

```

## RESULT 2

```

      052142 PRELIMINARY; PRT; 449 AA.
      AC 052142; 008194;
      DT 01-NOV-1996 (TReMBLrel. 01, Created)
      DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
      DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
      GN Iron-sulfur protein large subunit.
      OS Pseudomonas putida.
      OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
      NC NCB1_TaxID=303;
      RN [1]
      RP SEQUENCE FROM N.A.
      RC STRAIN=OUS82;
      RA Takizawa N.;
      RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
      RN [2]
      RP SEQUENCE FROM N.A.
      RC STRAIN=OUS82;
      RX MEDLINE=94209249; Pubmed=8157615;
      RA TAKIZAWA N., Kaide N., Torigoe S., Moritani T., Sawada T., Satoh S.,
      RA Kiyohara H.;
      RT "Identification and characterization of genes encoding polycyclic
      RT aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon
      RT dihydrodiol dehydrogenase in Pseudomonas putida OUS82.";
      RL J. Bacteriol. 176:2444-2449(1994).
      RN [3]
      RP SEQUENCE FROM N.A.
      RC STRAIN=OUS82;
      RA Noboru T., Toshiya I., Takashi S., Kazuhito Y., Masamichi K.,
      RA Yue-Wu W., Masao F., Hozoh K.;
      RT "The molecular analysis of NAH7-type cluster located on the
      RT chromosomes of Pseudomonas aeruginosa Pak1 and Pseudomonas putida
      RT OUS82.";
      RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
      DR EMBL: AB004059; BAA20391.1; -.
      DR HSSP: P23094; INDO.
      DR InterPro: IPR001281; Rieseke.
      DR InterPro: IPR001663; Ring_hydroxyl_A.
      DR Pfam: PF00848; Ring_hydroxyl_A.
      DR Pfam: PF00848; Ring_hydroxyl_A.
      DR PRINTS: PR00090; RINGDIHXGNASE.
      DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
      SQ SEQUENCE 449 AA: 49361 MW: 80020F54AB11E8A CRC64;

```

Query Match 95.0%; Score 2289; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 2; 8e-168;

```

      Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;
      QY 1 MNYNKILVSESGLSQKHILHGEDELFOHELKTIFFARNWMLFLTHDSLIIPAPGDYVAKMG 60
      1 MNYNKILVSESGLSQKHILHGEDELFOHELKTIFFARNWMLFLTHDSLIIPAPGDYVAKMG 60
      61 IDEVIVSRQNDGSIIRAFNLVCRHGRKTLVSVYAGNAKGFVCSYHGWGFGSNGELQSVPE 120
      61 IDEVIVSRQNDGSIIRAFNLVCRHGRKTLVSVYAGNAKGFVCSYHGWGFGSNGELQSVPE 120
      121 KDIYGESLNKRCGLGLKEVARSFHGFYCGFDEAPSLMDYLGDAAYLPIFKHSGL 180
      121 KELYGESLNKCLGLKEVARSFHGFYCGFDEAPSLMDYLGDAAYLPIFKHSGL 180
      181 ELVGPPKVVYIKAMWKAPEAFNFVGDAHYVGMTTHASSLRSGESIFSSLAGNALPPREGAGL 240
      181 ELVGPPKVVYIKAMWKAPEAFNFVGDAHYVGMTTHASSLRSGESIFSSLAGNALPPREGAGL 240
      241 QMTSKYSGMGVLWDGYSVHSADLVPELMAFGAKOERLNKEIGDVARIRYRSHLNCV 300
      241 QMTSKYSGMGVLWDGYSVHSADLVPELMAFGAKOERLNKEIGDVARIRYRSHLNCV 300
      301 FPNNSVLTGSGVFKVWNPIDANTTEVNTYALVEKDMPEDLKRLADAVQRTFGPAGFWS 360
      301 FPNNSVLTGSGVFKVWNPIDANTTEVNTYALVEKDMPEDLKRLADAVQRTFGPAGFWS 360
      361 DDNDNMTASONGKRYKSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
      361 DDNDNMTASONGKRYKSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
      421 QAHVSSSNMAEFEDASSTWHTELTKTTDR 449
      421 QAHVSSSNMAEFEDASSTWHTELTKTTDR 449

```

## RESULT 3

```

      092173 PRELIMINARY; PRT; 449 AA.
      AC 092173;
      DT 01-MAY-1999 (TReMBLrel. 10, Created)
      DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
      DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
      DE Naphthalene dioxygenase Fe-S large subunit.
      GN NAHAC.
      OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
      OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
      NC NCB1_TaxID=316;
      RN [1]
      RP SEQUENCE FROM N.A.
      RC STRAIN=AN10;
      RX MEDLINE=99365311; Pubmed=10433976;
      RA Bosch R., Garcia-Valdes E., Moore E.R.B.;
      RT "Genetic characterization and evolutionary implications of a
      RT chromosomally encoded naphthalene-degradation upper pathway from
      RT Pseudomonas stutzeri AN10.";
      RL Gene 236:149-157(1999).
      DR EMBL: AF039533; AAD02136.1; -.
      DR HSSP: P23094; INDO.
      DR InterPro: IPR001281; Rieseke.
      DR InterPro: IPR001663; Ring_hydroxyl_A.
      DR Pfam: PF00355; Rieseke; 1.
      DR Pfam: PF00848; Ring_hydroxyl_A.
      DR PRINTS: PR00090; RINGDIHXGNASE.
      DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
      KW Dioxygenase.
      SQ SEQUENCE 449 AA: 49802 MW: BA510FD1B6F66663 CRC64;

```

Query Match 91.5%; Score 2205; DB 2; Length 449;  
Best Local Similarity 89.5%; Pred. No. 1; 4e-161;  
Matches 402; Conservative 25; Mismatches 22; Indels 0; Gaps 0;

QY 1 MNYNKILVSESGLSQKHILHGEDELFOHELKTIFFARNWMLFLTHDSLIIPAPGDYVAKMG 60

Db	1	MMYKKNLTIVTSSGGLTQKHLIHGDEBELFQRELEFTTFARNMFLPHHDSLILBSPGDYVTAKKM	60
Qy	61	IDEIVYSHQNDGSIRAFELNVCRRHGKTLVSYEAGNAKGFVCSYHGMGFGSNGELQSVPE	120
Db	61	VDEVYVSHQNDGSIRAFELNVCRRHGKTLVHAEBGNAKGFVCSYHGMGFGANGLQSVPE	120
Qy	121	KDLIGESLNKCKCLGLEKVARYESHGFITYCCFQDEAPPLMDYIGDAAWYLEPFKHSGL	180
Db	121	KELYGALDKCKMGLKEVARYESHGFITYCCFDEEAPSLKDYMGDAGWYEPFKRHSGL	180
Qy	181	ELVYPPGVVYIKAMWKAPEFVGDATVHCVTHASSLRSESLFSSLAGNAALPEPGAGL	240
Db	181	ELIGPPGVYIIKAMWKAPEFVGDATVHCVTHASSLRTQSVETSLAGNAALPEPGAGL	240
Qy	241	QMTSYSGGMVLVDGYSGVHSAVLPELMAFGAKOERLKEIGVRAIRYSHLNCVY	300
Db	241	QMTSYSGGMVLVDGYSGVHSAVLPELMAFGAKOERLKEIGVRAIRYSHLNGVY	300
Qy	301	FPNNSMLTCSGVFEKVPIDANTTEVNTVAIVEKDMEDEDLKRLADSVORTGPAGFWES	360
Db	301	FPNNSFLTCSGVFEKVPIDANTTEVNTYAMVEKDMEDEDLKRLVDAYQOTEPAPFWES	360
Qy	361	DDNDMMETASONGKKYKOSRSDLSLNGFGEDEVYGAUVYPGVYKGAIGETSYRGFYAY	420
Db	361	DDNDMMETESONAKKYOSRQDGLSLNGFGEDEVYGVYVYKSAIGETSYRGFYAY	420
Qy	421	QAHYSSSMMAFEHJASSTWHTELTKTTDR	449
Db	421	GAHJSSSMMAFEHJASSTWHTELTKTTDR	449

Query Match	Similarity	Score	DB	Length
Best	90.9%	2190	DB 2	447
Matches	399	Conservative	27	Mismatches 21; Indels 2; Gaps 1

Query	Match	Similarity	Score	DB	Length
QY	1	MNYNNKILVSESGLSQKLLIHQDELFQHEKLTIPARWMLFTTHDSLTPAPGDYTYAKWG	60		
Db	1	M1YEN-LVSEAGLQKLLIHQDKRLFQHEKLTIPARWMLFTTHDSLTPAPGDYTYAKWG	58		
QY	61	IDEVYVSNQDNGSIRAPLVNCRHRKGLTVSVAGNAKGFVCSYHGWFQSGNELQSVPE	120		

Dd	59	VDVIVSNQNDGSVRAPLNTVORHNGKTLVHAENAGNAGFCVSHGWFSGSNGELQSVPE	118
Qy	121	KDLGESLNKKCLGIEKVARVESFHGFIYGCEDQADPLMDYLGDAAMYLEPMFKHSGT	180
Dd	119	KELYGDAIKKKCKLGIEKVPRIESHGFIYGCEDQADPLIDYLGDAAMYLEPIFKHSGGL	178
Qy	181	ELYGPCKVYVKAKMKKPAENFVDATVHGVTTHASSLRSGESIFSSLAGNAALPPBAGL	240
Dd	179	ELYGPCKVYVKAKMKKPAENFVDADAVHGVTTHASSLRSGOSIFTPPLAGNAMPPEBAGL	238
Qy	241	QMTSKYSGGMVLMDGASGVHSAADLVPELMAFGGAKOKEPINKETIGDVARITYRSHLCTV	300
Dd	239	QMTSKYSGGMVLMDAYSCHISADLVPEPMMAFGGAKOKEKLAKETIGDVARITYRSHLCTV	298
Qy	301	FPNNSMLTCSGVFKWPNIDANTDEVTYALVEKDMEDDLKRLRADSQRTGCPAGWES	360
Dd	299	FPNNSILTCGVFKWPNIDENITEVMTYALVEKDMEDDLKRLRADAVQRTGCPAGWES	358
Qy	361	DDNDMMETASONGKKKYOSRSDLSLNGFGEDVYGDVYGVGVGKSAIGETSYRGFYRAY	420
Dd	359	DDNDMMETESONAKKYOSSNDLIANTLNGFGEDVYGDCEYGVGVGKSAIGETSYRGFYRAY	418
Qy	421	QAHVSSNMWAEFEHASSTWHTELKTTDR	449
Dd	419	QAHVSSNMWAEFENTSRWHTELKTTDR	447

	RESULT 5			
ID	052382	PRELIMINARY:	PRT;	447 AA.
AC	052382			
DT	01-JUN-1998	(TREMBLrel. 06, Created)		
DT	01-JUN-1998	(TREMBLrel. 06, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	Naphthalene 1,2 dioxygenase large oxygenase component.			
GN	NgAcC.			
OS	Ralstonia sp. U2.			
OG	Plasmid pMWU2.			
OC	Bacteria; Proteobacteria; beta subdivision; Ralstonia group;			
CC	Ralstonia			
CX	NCBI_TaxID=70356;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-U2;			
RX	MEDLINE=98233751; PubMed=9573207;			
RA	Fuenmayor S.L., Wild M., Boyes A.L., Williams P.A.;			
RT	"A gene cluster encoding steps in conversion of naphthalene to			
RL	gentisate in pseudomonas sp. strain U2.";			
RN	J. Bacteriol. 180:2522-2530(1998).			
RP	[2]			
RC	SEQUENCE FROM N.A.			
RX	STRAIN=U2;			
RA	MEDLINE=20576173; PubMed=11133965;			
RT	Zhou N.Y., Fuenmayor S.L., Williams P.A.;			
RL	"nag genes of ralstonia (Formerly pseudomonas) sp. Strain U2 encoding			
DR	enzymes for gentisate catabolism.";			
DR	J. Bacteriol. 183:700-708(2001).			
DR	EMBL: AF036940; AAD12610.1; --			
DR	HSSP; P23094; INDO.			
DR	InterPro; IPR001281; Rieske.			
DR	InterPro; IPR001663; Ring_hydroxyl_A.			
DR	Pfam; PF00355; Rieske; 1.			
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.			
DR	PRINTS; PR00090; RNCDIOXNAS.			
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.			
KW	Dioxygenase; Plasmid.			
SQ	SEQUENCE 447 AA; 49570 MW; 4553AAFAB4410EDO CRC64;			
	Query Match	90.7%; Score 2186; DB 2; Length 447;		
	Best local Similarity	89.1%; Pred. No. 4.1e-160;		
	Matches 400; Conservative 25; Mismatches 22; Indels 2; Gaps 1;			
Y	1 MNYNKRIIVSESGLSQKHIIHODELPFCHKEKTIPARMMILFTTHDSLIPAPGVYTAKKG 60			

```
Db 1 MYEN--LVSEAGLTQKHLIHGDKELFQHHELTIFARNWMLFLTHDSLIPSPGDVYAKMG 58
Qy 61 IDEVIYSRONDGSTRALFVNCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
Db 59 VDEVIYSRONDGSTRALFVNCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 118
Qy 121 KDLVGESENKRCCLGKLEKVARVESFHGFTYGCDFDAPPLMDYLGDAAMYLEPMFKHSGGL 180
Db 119 KELXGDTIKKKCLGKLEKVARVESFHGFTYGCDFDAPPLMDYLGDAAMYLEPMFKHSGGL 178
Qy 181 ELVGPCKVVIKANKWKAPEENFVGDAYHVGWTHASSLRSESIFFSLAGNALPPEGAGL 240
Db 179 ELVGPCKVVIKANKWKAPEENFVGDAYHVGWTHASSLRSESIFFSLAGNALPPEGAGL 238
Qy 241 QMTSKYSGMGVLMDGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYSHLNTCV 300
Db 239 QMTSKYSGMGVLMDGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYSHLNTCV 298
Qy 301 FPNNSMLTCSGVFVKNPDPIDANTEVNTYAIVEKMPEDLKRRLADSVORTGPAGFWES 360
Db 299 FPNNSMLTCSGVFVKNPDPIDANTEVNTYAIVEKMPEDLKRRLADSVORTGPAGFWES 358
Qy 361 DDNDNMTASONGKRYOSRSDLLSNLGFGEDEVYGDVAVPGVKSALGETSYRGFYRAY 420
Db 359 DDNDNMTESQNAKKYOSNSDLIANLGFGEDEVYGDVAVPGVKSALGETSYRGFYRAY 418
Qy 421 QAHVSSNMAFEHASSTWHTELTKTTDR 449
Db 419 QAHVSSNMAFEHASSTWHTELTKTTDR 447

RESULT 6
Q8VUD4 PRELIMINARY: PRT: 447 AA.
ID 08VUD4;
AC 08VUD4;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Dntac.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group.
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-R34.
RX MEDLINE=20254695; PubMed=10795678;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Properties of the trihydroxytoluene oxygenase from Burkholderia
   cepacia R34: an extradiol dioxygenase from the 2,4-dinitrotoluene
   pathway.";
RL Arch. Microbiol. 173:86-90(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-R34.
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Origins of the 2,4-dinitrotoluene pathway.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF169302; AAL50021.1; -.
DR InterPro; IPR001281; Rieseke.
DR Pfam; PF00355; Rieseke.1.
DR Pfam; PF00848; Ring_hydroxyl_A.1.
DR PRINTS; PR00090; RINGDIPOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.
SQ SEQUENCE 447 AA; 49468 MW; 2A5BF8558320275C CRC64;

Query Match 85.8%; Score 2068; DB 2; Length 447;
Best Local Similarity 83.5%; Pred. No. 5e-151;
Matches 375; Conservative 38; Mismatches 34; Indels 2; Gaps 1;
```

```
Db 1 MSYON--LVSEAGLTQKHLIHGDKELFQHHELTIFARNWMLFLTHDSLIPSPGDVYAKMG 58
Qy 61 IDEVIYSRONDGSTRALFVNCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
Db 59 VDEVIYSRONDGSTRALFVNCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 118
Qy 121 KDLVGESENKRCCLGKLEKVARVESFHGFTYGCDFDAPPLMDYLGDAAMYLEPMFKHSGGL 180
Db 119 KELXGDTIKKKCLGKLEKVARVESFHGFTYGCDFDAPPLMDYLGDAAMYLEPMFKHSGGL 178
Qy 181 ELVGPCKVVIKANKWKAPEENFVGDAYHVGWTHASSLRSESIFFSLAGNALPPEGAGL 240
Db 179 ELVGPCKVVIKANKWKAPEENFVGDAYHVGWTHASSLRSESIFFSLAGNALPPEGAGL 238
Qy 241 QMTSKYSGMGVLMDGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYSHLNTCV 300
Db 239 QMTSKYSGMGVLMDGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYSHLNTCV 298
Qy 301 FPNNSMLTCSGVFVKNPDPIDANTEVNTYAIVEKMPEDLKRRLADSVORTGPAGFWES 360
Db 299 FPNNSMLTCSGVFVKNPDPIDANTEVNTYAIVEKMPEDLKRRLADSVORTGPAGFWES 358
Qy 361 DDNDNMTASONGKRYOSRSDLLSNLGFGEDEVYGDVAVPGVKSALGETSYRGFYRAY 420
Db 359 DDNDNMTLSQNAKKYOSNSDLIADLGFGEDEVYGDVAVPGVKSALGETSYRGFYRAY 418
Qy 421 QAHVSSNMAFEHASSTWHTELTKTTDR 449
Db 419 QAHVSSNMAFEHASSTWHTELTKTTDR 447

RESULT 7
P95564 PRELIMINARY: PRT: 447 AA.
ID P95564;
AC P95564;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Ispalpa 2NT.
CN NPDAC.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JS42.
RX MEDLINE=97128768; PubMed=8973308;
RA Parales J.V., Kumar A., Parales R.E., Gibson D.T.;
RT "Cloning and sequencing of the genes encoding 2-nitrotoluene
   dioxygenase from Pseudomonas sp. JS42.";
RL Gene 181:57-61(1996).
DR EMBL; U49504; AAB40383.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieseke.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieseke.1.
DR Pfam; PF00848; Ring_hydroxyl_A.1.
DR PRINTS; PR00090; RINGDIPOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 447 AA; 49485 MW; ICB0E223E528E3BD CRC64;

Query Match 85.0%; Score 2048; DB 2; Length 447;
Best Local Similarity 83.7%; Pred. No. 1.7e-149;
Matches 376; Conservative 34; Mismatches 37; Indels 2; Gaps 1;
```

```
Qy 1 MNYNKLIVSESGLSQKHLIHGDELFQHELTIFARNWMLFLTHDSLIPAPGDVYAKMG 60
Db 1 MSYON--LVSEAGLTQKHLIHGDKELFQHHELTIFARNWMLFLTHDSLIPSPGDVYAKMG 58
Qy 61 IDEVIYSRONDGSTRALFVNCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
Db 59 VDEVIYSRONDGSTRALFVNCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 118
```





```
Db      ||||| ||||| :|||||:||||| ||||| || :|||:|||||
364 DDNDNM-VLSGNAKKYYQSSNDLADJGFGDYIGDCECPGVSKSAFSETHNGFTRAY 42222
Oy      421 QAHVSSNNAAEFEEHASTFWHTELTKTTDR 449
        |||:|||||: | |||||:|
423 QAHISSNNAAFEENTSRNWHTELTCTTDR 451
```

## RESULT 10

ID	Q9ETK2	PRELIMINARY;	PRT;	277	AA
----	--------	--------------	------	-----	----

DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment)

OS *Pseudomonas putida*.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.  
OC *Pseudomonas*.  
OX NCBI\_TaxID=303;

RP SEQUENCE FROM N. A.  
RC STRAIN=2IDINH, PRIM1, AND 3IA2NH;  
RA Ferrero M.A., Lalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahAc) in  
PL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DT 01-MAR-2001 (TREMblref. 16, last sequence update)  
 DT 01-DEC-2001 (TREMblref. 19, last annotation update)  
 DE Naphthalene dioxygenase iron sulfur protein 2 (fragment).  
 GN NAHAC2.  
 OS Pseudomonas putida.  
 OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OC NCBI\_taxid=303;

RP SEQUENCE FROM N.A.  
RC STRAIN=5111ASAL:  
RA Ferrero M.A., Taluati J., Bosch R.;  
RT \*Coexistence of two naphthalene dioxygenase genes (nahac) in  
Pseudomonas strains from West Mediterranean Sea.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

ID	09F554	PRELIMINARY;	PRT;	277	AA.	RT	"Coexistence of two naphthalene dioxygenase genes (nahac) in
AC	09F554;					RT	Pseudomonas strains from West Mediterranean Sea.";
DT	01-MAR-2001	(TREMBLrel. 16, Created)				RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR	EMBL: AF306426; AAC25686.1; -
DR	HSSP: P23094; INDO.
DR	InterPro: IPR001281; Rieske.
DR	InterPro: IPR001663; Ring_hydroxyl_A.
DR	Pfam: PF00355; Rieske; 1.
DR	Pfam: PF00848; Ring_hydroxyl_A; 1.
DR	PRINTS: PR00090; RINGDIOXGNASE.
DR	PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW	Dioxygenase.
FT	NON_TER . 1 1
FT	NON_TER 277 277
Q	SEQUENCE 277 AA; 30064 MW; DB47868EB6D525A CRC64;

Query Match	61.7%	Score 1486;	DB 2;	Length 277;
Best Local Similarity	99.6%	Pred. No. 1.7e-106;		
Matches 276; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	56	1	60
TAKMGIDEIVSRÖNDSDIRAFI	TAKMGIDEIVSRÖNDSDIRAFI	TAKMGIDEIVSRÖNDSDIRAFI	TAKMGIDEIVSRÖNDSDIRAFI
NCVCRHGRKTLVSVEGNNAKFCV	NCVCRHGRKTLVSVEGNNAKFCV	NCVCRHGRKTLVSVEGNNAKFCV	NCVCRHGRKTLVSVEGNNAKFCV
SYTHGWSGSGNGEIQ	SYTHGWSGSGNGEIQ	SYTHGWSGSGNGEIQ	SYTHGWSGSGNGEIQ
11.5	11.5	11.5	11.5

Qy 116 SVPEEDDLGEGELNKKKGLKEAVARVSEFHGFTYGCFOEAPRLMDYGDAAVLEPER 175  
Db 61 GVPFEKDLGEGELNKKKGLKEAVARVSEFHGFTYGCFOEAPRLMDYGDAAVLEPER 120

Qy	176	HSGGLELVGPEKVKVIAKMKRPAENFVGDAHVGMTHASSLSGSESTFSSLSLGNALPP	235
Db	121	HSGGLELVGPEKVKVIAKMKRPAENFVGDAHVGMTHASSLSGSESTFSSLSLGNALPP	180

QY	236	EGAGLQMTSKTSGMGVLTMDGYSGVHSADIVPELMAKCGAKORLNKEIGDVARITYSSH	295
Db	181	EGAGLQMTSKTSGMGVLTMDGYSGVHSADIVPELMAKCGAKORLNKEIGDVARITYSSH	240

Qy	296	LNCTVEPNNSMLTCSGYFKVWNPIDANTTEVMTYAIV	332
Db	241	LNCTVEPNNSMLTCSGYFKVWNPIDANTTEVMTYAIV	277

RESULT 13  
09F5S8

ID	Q9F5S8	PRELIMINARY;	PRT;	277 AA.
AC	Q9F5S8			
DT	01-MAR-2001	(TREMBLrel. 16, Created)		
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)		
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)		
DE	Naphthalene dioxygenase iron sulfur protein (Fragment).			

CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Pseudomonas  
 CC NCBI\_TaxID=139773;

```

KN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=PR3MN2;

```

RT "Coexistence of two naphthalene dioxygenase genes (nahAc) in  
RT Pseudomonas strains from West Mediterranean Sea.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR HSSP; P23034; INDO.  
DR InterPro; IPR001281; Rieske.  
DR InterPro; IPR001663; Ring\_hydroxyl\_A.

```
DR pfam: PF00848; RING_HYDROXYL_A; 1.
DR PRINTS: PR00090; RINGDIPOXINASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
```

FT	1	1	
NON_TER	277	277	
FT	277	277	
SEQUENCE	277	277	
AA:	30064	MM:	6BFF83D32F04CFC3
CRC64:			

Query Match	61.68;	Score 1484;	DB 2;	Length 277;
Best Local Similarity	99.68;	Pred. No. 2.4e-106;		

Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

56 TAKMGIDEVLSRQNDGSIKRAFLNVCRRHGKTLVSVEAGNAKGFCSYHMGFGSGNELQ 115  
b 1 TAKMGIDEVLSRQNDGSIKRAFLNVCRRHGKTLVSVEAGNAKGFCSYHMGFGSGNELQ 60

b y  
116 SVPFEDLYGESLNKKCLGLKEARVESFHGFIYGCFDQAPRLMDYLGDAAWYLEPMFK 175  
|||||  
61 SVPFEDLYGESLNKKCLGLKEARVESFHGFIYGCFDQAPRLMDYLGDAAWYLEPMFK 120

176 HSGGELVGPCKVVIKANMKAPAENFVGDAYHHGWTTHASSLRSESIFFSLAGNALPP 235  
y  
121 HSGGELVGPCKVVIKANMKAPAENFVGDAYHHGWTTHASSLRSESIFFSLAGNALPP 180  
b

y 236 EGAGDQMTSKGSGMGVLMDGSGVHSADLPPELMAFGAKQERLNKEIGVRRARYRSH 295  
 |||||  
 b 181 EGAGIQTTSKGSGMGVLMDGSGVHSADLPPELMAFGAKQERLNKEIGVRRARYRSH 240

[illegible]

RESULT 14  
9F5S9

C Q9F5S9;  
T 01-MAR-2001 (TREMBLrel, 16, Created)  
T 01-MAR-2001 (TREMBLrel, 16, Last sequence update)  
01-MAR-2001 (TREMBLrel, 16, Last sequence update)

E Naphthalene dioxygenase iron sulfur protein (Fragment).  
 S NAHAC.  
 N Pseudomonas sp. 81DINH.

C Pseudomonas.  
X NCBI\_TaxID=139736;  
N [1]  
N SEQUENCE FROM N A

STRAIN=8IDINH;  
C  
Ferrero M.A., Talucat J., Bosch R.;  
A  
"Coexistence of two naphthalene dioxygenase genes (nahc) in  
T  
broadening studies from host modification of the "

Submitted (Sep-2000) to the EMBL/Genbank/DBJ databases.  
 R EMBL: AF306433; AAG25693.1; -  
 R HSSP: P23094; INDO.  
 R TatCoDys: TDB001281; P15036

R InterPro; PR001663; Ring\_hydroxyl\_A.  
R Pfam: PF00355; Rieske; 1.  
R Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DRMS: DR00000. DMSOVSCHEN

R	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
W	dioxygenase.
T	NON_TER 1
N	NON_TER 327

Q	SEQUENCE	277 AA;	29995 MW;	AADAGB56378FF6DB CRC64;
Query Match	61.5%	Score 1483;	DB 2;	Length 277;
Best Local Alignment	60.6%	Score 1483	DB 2	Length 277

Matches 276, Conservative 0; Mismatches 1; Indels 0; Gaps 0;

116 SVPFEDLYGSLNKKCLGLKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYIEPMFK 175

176 HSGGLELVGPBGKVVITANKKAPAENEYGDAYHYGWTHTASSLRSGESTIPSSLAGNALPP 235

236 EGAGLQMTSKYGGSGMGLWMDISGVHSADLVPELMAFGAKQERLNKETGDVRRARIYRSH 295

Db 181 EGAGLQMTSKYSGWGLMDYSGVSHADLVPBLMAFGAKOGLNKELGIVRARIYRSH 240  
QY 296 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 332  
Db 241 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 277

## RESULT 15

Q9F5S3 PRELIMINARY; PRT: 277 AA.  
AC Q9F5S3;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 1 (Fragment).  
GN NHAHL.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ST1ANH;  
RA Ferrero M.A., Jalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahA) in  
Pseudomonas strains from West Mediterranean Sea."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF306441; AAC25701.1; .  
DR HSP; P23094; INDO.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIOXNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
FT NON\_TER 1 1  
FT 277 277  
SQ SEQUENCE 277 AA; 30062 MW; 82C78D865A92D0AB CRC64;

Query Match 60.2%; Score 1451; DB 2; Length 277;  
Best Local Similarity 96.8%; Pred. No. 8.3e-104;  
Matches 268; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 56 TAKKGIDEVITYSRNDGSIKRAFLNCRHRGKTIYVEAGNAKGFYCSYHGMGFGSNGELQ 115  
Db 1 TAKKGIDEVITYSRDSDGSIKRAFLNCRHRGKTLVNAEAGNAKGFVCSYHGMGFGSNGELQ 60  
QY 116 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCFFDQEAAPLMDYLGDAAMYLEPMFK 175  
Db 61 SVPEKELYGESLNKKCLGLKEVARVESFHGFIYGCFFDQEAAPLMDYLGDAAMYLEPIFK 120  
QY 176 HSGGLELVGPPGKVIYKANMKAPAEVGVDAVHVGWTHASSIRSGESIFSSLAGNAALPP 235  
Db 121 HSGGLELVGPPGKVIYKANMKAPAEVGVDAVHVGWTHASSIRSGESIFSSLAGNAALPP 180  
QY 236 EGAGLQMTSKYSGWGLMDYSGVSHADLVPBLMAFGAKOERLNKEIGVRAIRYRSH 295  
Db 181 EGAGLQMTSKYSGWGLMDYSGVSHADLVPBLMAFGAKOERLNKEIGVRAIRYRSH 240  
QY 296 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 332  
Db 241 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 277

Search completed: January 27, 2003, 09:02:19  
Job time : 27.8333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:47:32 ; Search time 7.16667 Seconds  
(without alignments)  
2598.540 Million cell updates/sec

Title: US-09-843-250-32

Sequence: 1 MNVNNKILVSESGLSOKHLI.....AEFEHASSVWHELTAKTDR 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2401	99.6	449	1 NDOB_PSEPU	P23094 pseudomonas
2	2315	96.1	449	1 NDOB_PSEFL	O07824 pseudomonas
3	2216	92.0	449	1 NDOB_PSEAE	O51494 pseudomonas
4	734	30.5	453	1 HCAE_ECOLI	O47139 escherichia
5	649.5	27.0	450	1 BED1_PSEPU	O07844 pseudomonas
6	647.5	26.9	450	1 TOD1_PSEPU	P13450 pseudomonas
7	614.5	25.5	458	1 BPHA_BURCE	P37333 burkholderi
8	614	25.5	457	1 BPHA_COMTE	O46372 comamonas t
9	614	25.5	458	1 BPHA_PSEPS	P08084 pseudomonas
10	586.5	24.3	448	1 BNZA_PSEPU	O52028 pseudomonas
11	584	24.2	458	1 BPA1_PSESL	O52438 pseudomonas
12	410.5	17.0	454	1 XYLK_PSEPU	P23099 pseudomonas
13	396	16.4	461	1 BEVA_ACICA	P07769 acinetobact
14	244	10.1	374	1 YEAM_ECOLI	P76253 escherichia
15	239.5	9.9	442	1 CHMO_AMATR	O93x01 amarantus
16	233	9.7	438	1 CHMO_ATRHO	O91x00 atriplex ho
17	224	9.3	446	1 CHMO_BETVU	O22553 beta vulgar
18	223	9.3	439	1 CHMO_SPTOL	O04121 spiruacia ol
19	216	9.0	422	1 CHMO_APATH	O95x00 arabidopsis
20	165	6.8	439	1 PHT3_PSEPU	O05183 pseudomonas
21	118.5	4.9	354	1 VANA_PSESP	O05616 pseudomonas
22	113	4.7	402	1 CBAA_COMTE	O44256 comamonas t
23	112	4.6	409	1 POBA_PSEPS	O52185 pseudomonas
24	110	4.6	329	1 VANA_PSEPS	P12609 pseudomonas
25	106.5	4.4	1926	1 LPH_RABIT	P09849 oryctolagus
26	106	4.4	1276	1 PMP6_CHLPN	O92899 chlamydia p
27	101.5	4.2	3255	1 POLG_LMVE	P89876 1 genome po
28	100.5	4.2	543	1 TUP1_DROME	P16375 drosophila
29	100.5	4.2	746	1 TUP2_DROME	P16376 drosophila
30	100.5	4.2	1411	1 Y297_HUMAN	O15040 homo sapien
31	97.5	4.0	3255	1 POLG_LMVO	P31999 1 genome po
32	94.5	3.9	331	1 LDHA_RHIDE	O9p58 rhizophila
33	94.5	3.9	350	1 UNRI_HUMAN	O9y3f4 homo sapien

34	94.5	3.9	452	1 F26_YEAST	P32604 saccharomyc
35	94.5	3.9	468	1 PPA1_PICPA	P52291 picchia past
36	94.5	3.9	518	1 ATPA_ENTHR	P26679 enterococcu
37	94	3.9	847	1 ORB8_HUMAN	O9b2f1 homo sapien
38	93	3.9	420	1 DHE3_PYROO	O59650 pyrococcus
39	93	3.9	762	1 SLAP_ACEKI	P22258 acetogenium
40	93	3.9	2261	1 ABC1_MOUSE	P41233 mus musculu
41	92	3.8	985	1 AGU1_ASPOR	O12558 aspergillus
42	92	3.8	1693	1 POLN_HEVMY	O04610 hepatitis e
43	91.5	3.8	272	1 SER3_DROME	P17207 drosophila
44	91.5	3.8	525	1 NCAP_MEASA	P35972 measles vir
45	91	3.8	917	1 SYL_STAAU	P41972 staphylococ

## ALIGNMENTS

RESULT 1  
ID NDOB\_PSEPU STANDARD: PRT: 449 AA.  
AC P23094; O52124; O33461; O07830;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene 1,2-dioxygenase ISP alpha).  
GN NDOB OR NAHAC OR DOXB OR NAHA3 OR NDOC2.  
OS Pseudomonas putida,  
OS Pseudomonas sp. (strain C18).  
OG Plasmid pDTG1, Plasmid NAH7, and Plasmid NPL1.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303, 306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=89211973; PubMed=3243438;  
RA "Kurela S., Leivaeslahti H., Palva E.T., Teerl T.H.;  
RT "Cloning, nucleotide sequence and characterization of genes encoding  
RT naphthalene dioxygenase of Pseudomonas putida strain NCIB9816.";  
RL Gene 73:355-362(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX Parales J.V., Parales R.E., Kumar A., Gibson D.T.;  
RT Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C18;  
RX MEDLINE=94042852; PubMed=8226631;  
RA Denome S.A., Stanley D.C., Olson E.S., Young K.D.;  
RT "Metabolism of dibenzothophene and naphthalene in Pseudomonas  
RT strains: complete DNA sequence of an upper naphthalene catabolic  
RT pathway.";  
RL J. Bacteriol. 175:6890-6901(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=G7 / ATCC 17485; PLASMID=NAH7;  
RX MEDLINE=93522277; PubMed=8486285;  
RA Simon M.J., Oselund T.D., Saunders R., Ensley B.D., Suggs S.,  
RT Harcourt A.A., Sun W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;  
RT "Sequences of genes encoding naphthalene dioxygenase in Pseudomonas  
RT putida strains G7 and NCIB 9816-4.";  
RL Gene 127:31-37(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=BS202; PLASMID=NPL1;  
RX Bezborodnikov S.G., Boronin A.M., Tiedje J.M.;  
RT "Nucleotide sequences of genes encoding an upper pathway of  
RT naphthalene metabolism of NPL1 plasmid from Pseudomonas putida strain  
RT BS202.";  
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.  
RN [6]

```

RP SEQUENCE FROM N.A. STRAIN-ATCC 17484;
RC SPECIES=P.putida;
RA Hamann C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).
RC SPECIES=P.putida; STRAIN=NCIB 9816;
RX MEDLINE=96298434; PubMed=96346935;
RA Knappl B., Lee K., Carriedano E., Perales R.E., Gibson D.T., Eklund H.,
RA Ramaswamy S.;
RT "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene
RT 1,2-dioxygenase.";
RL Structure 6:571-586(1998).
CC -I- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL.
CC -I- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -I- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -I- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIENANTHROPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -I- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF AN HEXAMER OF THREE LARGE ALPHA SUBUNITS (NDOB) AND
CC THREE SMALL BETA SUBUNITS (NDOC).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/sib.ch).
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M23914; AAB47591.1; -
DR EMBL; U49496; AAA92141.1; -
DR EMBL; M60405; AAA16125.1; -
DR EMBL; M83949; AAA25902.1; -
DR EMBL; AF010471; AAB62707.1; -
DR EMBL; AF004284; AAB61373.1; -
DR PIR; JS0071; JS0071.
DR PIR; B49343; B49343.
DR PDB; 1NDO; 23-MAR-99.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PRO0090; RINGDIOXGNSE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD; Plasmid; 3d-structure.
FT METAL 81 81 IRON-SULFUR (2FE-2S).
FT METAL 83 83 IRON-SULFUR (2FE-2S).
FT METAL 101 101 IRON-SULFUR (2FE-2S).
FT METAL 104 104 IRON-SULFUR (2FE-2S).
FT METAL 208 208 IRON.
FT METAL 213 213 IRON.
FT METAL 362 362 IRON.
FT VARIANT 4 4 N -> K (IN STRAIN G7).
FT VARIANT 12 12 S -> F (IN STRAIN ATCC 17484).
FT VARIANT 15 15 S -> T (IN STRAIN G7).
FT VARIANT 32 32 K -> R (IN STRAIN G7).
FT VARIANT 50 50 A -> S (IN STRAIN G7).
FT VARIANT 70 70 N -> S (IN STRAIN G7).
FT VARIANT 90 91 SV -> NA (IN STRAIN G7).

```

Query Match	Best Local Similarity	Score 2401:	DB 1:	Length 449:
Matches 448:	Conservative	0:	Mismatches 1:	Indels 0: Gaps 0:
QY 1 MNYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARWMLFETHSLIPAPGDVYTAKMG 60	173	D -> E (IN STRAIN G7).		
DB 1 MNYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARWMLFETHSLIPAPGDVYTAKMG 60	173	M -> I (IN STRAIN G7).		
QY 61 IDEVIVSRONDGSIRAFELVNCRHRRKTLTVSFAAGNAKGFVCSYHGSGFSGNGELOSVPFE 120	225	S -> A (IN STRAIN G7).		
DB 61 IDEVIVSRONDGSIRAFELVNCRHRRKTLTVSFAAGNAKGFVCSYHGSGFSGNGELOSVPFE 120	225	S -> C (IN STRAIN BS202).		
QY 121 KOLYGSLSLKKKLGLEKAVARVESHFGFTYGCDDQAPPLMDYLGQAANYLERPMKRHSGL 180	232	A -> V (IN STRAIN G7).		
DB 121 KOLYGSLSLKKKLGLEKAVARVESHFGFTYGCDDQAPPLMDYLGQAANYLERPMKRHSGL 180	232	A -> S (IN STRAIN G7).		
QY 181 ELVGPFGKVIKANKKAPARENFGAVAHGWTTHASLSRGSIESFSLAGNALPPEGAGL 240	275	E -> K (IN STRAIN G7).		
DB 181 ELVGPFGKVIKANKKAPARENFGAVAHGWTTHASLSRGSIESFSLAGNALPPEGAGL 240	275	Q -> R (IN STRAIN ATCC 17484).		
QY 241 QMTSKYSGMGLMDGYSVSHSADLVPELMAFGAKOERLKEIGDVBARITYRSHLNTCV 300	391	H -> D (IN STRAIN G7C64).		
DB 241 QMTSKYSGMGLMDGYSVSHSADLVPELMAFGAKOERLKEIGDVBARITYRSHLNTCV 300	421			
QY 301 FPNNSMLTCSGYFKYWNPIDANTTEVWYIAYEKRMPEDLKRRLADSVQRTGGPAGFWES 360	434			
DB 301 FPNNSMLTCSGYFKYWNPIDANTTEVWYIAYEKRMPEDLKRRLADSVQRTGGPAGFWES 360	434			
QY 361 DONDMMETASQKKYQSHRSDLSNLGFGEDVDYGDAYPGVYGKSAIGETSYRGFTYAY 420	449			
DB 361 DONDMMETASQKKYQSHRSDLSNLGFGEDVDYGDAYPGVYGKSAIGETSYRGFTYAY 420	449			
QY 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449	449			
DB 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449	449			
RESULT 2				
NDOB_PSEFL	STANDARD:	PRT:	449 AA.	
AC 007824;				
DT 15-DEC-1998 (Rel. 37, Created)				
DT 15-DEC-1998 (Rel. 37, Last sequence update)				
DT 15-JUN-2002 (Rel. 41, Last annotation update)				
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene 1,2-dioxygenase ISP alpha).				
GN NDOB OR NDOC2.				
OS Pseudomonas fluorescens.				
OG Plasmid.				
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;				
OC Pseudomonas.				
OC NCBI_TaxID=294;				
RP SEQUENCE FROM N.A.				
RC STRAIN-ATCC 17483;				
RA Hamann C.;				
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.				
CC -I- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)				
CC MUTATION: COMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF				
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-				
CC NAPHTHALENE DIHYDRODIOL.				
CC -I- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-				

```
CC dihydronaphthalene-1,2-diol + NAD(+).
CC COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF004283; AAB61370.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske_1.
CC Pfam: PF00848; Ring_hydroxyl_A; 1.
CC PRINTS: PR00090; RINGDIOXGNASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC KM
CC DIOXYGENASE: NAD; Plasmid.
CC FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 208 208 IRON (BY SIMILARITY).
CC FT METAL 213 213 IRON (BY SIMILARITY).
CC FT METAL 362 362 IRON (BY SIMILARITY).
CC SEQIDCE 449 AA; 49512 MW; 186A7924AD437A1C CRC64;

Query Match 96.1%; Score 2315; DB 1; Length 449;
Best Local Similarity 95.5%; Pred. No. 2e-173;
Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 MNYNKTIIVSESGSLKHLIHGDELFQHELKTIFARNMLFTTHDSLIPAGDYVTAKKG 60
DB 1 MNYNKTIIVSESGSLKHLIHGDELFQHELKTIFARNMLFTTHDSLIPAGDYVTAKKG 60
CC
QY 61 IDEVIYSRONDSIRAFILVWCRRHGRKTLVSEAGNAKGFVCSYHGWFSGNGLQSVPEE 120
DB 61 IDEVIYSRONDSIRAFILVWCRRHGRKTLVSEAGNAKGFVCSYHGWFSGNGLQSVPEE 120
CC
QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCPEQAPLMDYLGDAAWLEPFKHSGL 180
DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCPEQAPLMDYLGDAAWLEPFKHSGL 180
CC
QY 121 KELYGSLLNKKCLGLKEVARVESFHGFIYGCPEQAPLMDYLGDAAWLEPFKHSGL 180
DB 121 KELYGSLLNKKCLGLKEVARVESFHGFIYGCPEQAPLMDYLGDAAWLEPFKHSGL 180
CC
QY 181 ELVGPFGKVIYKANKAPAPENYGDAYHWGTHASSLRSGESIFSSLAGNAALPPEGAGL 240
DB 181 ELVGPFGKVIYKANKAPAPENYGDAYHWGTHASSLRSGESIFSSLAGNAALPPEGAGL 240
CC
QY 241 QMTSKVSGSGVLMDSYSGVHSADLPELMARFGAEOERLNEIGVPRARIYRSHLNCV 300
DB 241 QMTSKVSGSGVLMDSYSGVHSADLPELMARFGAEOERLNEIGVPRARIYRSHLNCV 300
CC
QY 301 FPNNSMLTSGVGFVKWNPIDANTTEVWTAYAIYEKMPEDLKRRLADSVORTGSPAGFWS 360
DB 301 FPNNSMLTSGVGFVKWNPIDANTTEVWTAYAIYEKMPEDLKRRLADSVORTGSPAGFWS 360
CC
QY 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYGDAYVPGVVGKSAIGETSYRGFYRAY 420
DB 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYGDAYVPGVVGKSAIGETSYRGFYRAY 420
CC
```

```
DB 361 DDNDNMETASQNGKKYQSRDLSLNLGFGEDYGDAYVPGVVGKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNNAEPEFASSTWHTTEKTTDR 449
DB 421 QAHVSSNNAEPEFASSTWHTTEKTTDR 449
CC
RESULT 3
NDOB.PSEAE
ID NDOB.PSEAE STANDARD: PRT: 449 AA.
AC 05194;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DE 15-JUN-2002 (rel. 41, Last annotation update)
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene
DE 1,2-dioxygenase ISP alpha).
GN NDOB OR PAK43.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAK1;
RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,
RA Kiyohara H.;
RT "The molecular analysis of an NAH7-type gene cluster, pah, located on
RT the chromosome of Pseudomonas aeruginosa PAK1."
RL Submitted (Apr-1996) to the EMBL/Genbank/DDJ databases.
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL.
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D84146; BAA12240.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske_1.
CC Pfam: PF00848; Ring_hydroxyl_A; 1.
CC PRINTS: PR00090; RINGDIOXGNASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC KM
CC DIOXYGENASE: NAD.
CC FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 208 208 IRON (BY SIMILARITY).
```

```

FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 362 362 IRON (BY SIMILARITY).
SQ SEQUENCE 449 AA; 49715 MW; 35A189136722A21C CnC64;

Query Match 92.0%; Score 2216; DB 1; Length 449;
Best Local Similarity 89.8%; Pred. No. 1.1e-165;
Matches 403; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

OY 1 MNYNKKIVSSSGLSQKLLIHGDELLFQHELTIFARNMWLFTHSLPAPGDYATKMG 60
DB 1 MNYNKKIVSSSGLSQKLLIHGDELLFQHELTIFARNMWLFTHSLPAPGDYATKMG 60
OY 61 IDEVIVSRQNDGSIKRAFLNVCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
DB 61 VDEVIVSRQNDGSIKRAFLNVCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
OY 121 KDLYGESLNKCKLGEKAVARVESFHGFIYGCEDQAPPLMDYLGAANYLEPMFHSGL 180
DB 121 KELYGEALDKCKCMGKEKAVARVESFHGFIYGCEDQAPPLMDYGAANYLEPMFHSGL 180
OY 181 ELVGPGRKVVIRKANKKAPAEENVGDYHVGMTHASSLSRSGESIFSSLAGNALPPEGAGL 240
DB 181 ELIIPPGRKVIIRKANKKAPAEENVGDYHVGMTHASSLSRSGESIFSSLAGNALPPEGAGL 240
OY 241 QMTSKYSGMGVLMDGYSVGSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNTCTV 300
DB 241 QMTSKYSGMGVLMDGYSVGSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNTCTV 300
OY 301 FPNNSMLTCSGVFKVMNIDANTFEVWTYAYVEKMPEDLKRRLDSVQRTGPGAFWES 360
DB 301 FPNNSFLTCSGVFKVMHIDANTFEVWTYAYVEKMPEDLKRRLDAVQRTGPGAFWES 360
OY 361 DDNNMETASNGKKYGRSDLSNLGFGEDYGDVAYPGVWGSAVGETSYRFEYRAY 420
DB 361 DDNNMEIVSQNAKKYGRSDGLVNLGFGEDYGVDEYVPGVWGSAVGETSYRFEYRAY 420
OY 421 QAHVSSNMWAEFEHSASTWHTELTKTDR 449
DB 421 GAHISSSWAEFEEDVSKNMHTELAKTTDR 449

RESULT 4
HCAE_ECOLI STANDARD; PRT; 453 AA.
AC Q47139; P77590; P78203;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-phenylpropionate dioxygenase alpha subunit (EC 1.14.1.-)
GN HCAE OR PHD1 OR HCAA OR HCAAL OR DIGA OR B2538 OR Z3809 OR ECS3404.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_taxid=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RL MEDLINE=9742617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN-K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubraman S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP FUNCTION: PART OF MULTICOMPONENT 3-PHENYLPROPIONATE DIOXYGENASE.
CC CONVERTS 3-PHENYLPROPIONIC ACID (PP) INTO CIS-3-(3-CARBOXYETHYL)-
CC 3,5-CYCLOHEXADIENE-1,2-DIOL (PP-DIHYDRODIOL).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: 3-phenylpropionic acid catabolism.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (HCAE AND HCAF), A
CC FERREDOXIN (HCAC) AND A FERREDOXIN REDUCTASE (HCAD).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z37966; CAAB6018.1; -
CC EMBL: AE000340; AAC75591.1; -
CC EMBL: D90883; BAAL6433.1; -
CC EMBL: D90884; BAAL6441.1; -
CC EMBL: AE005484; AAC57651.1; -
CC EMBL: AP002562; BAB36827.1; -
CC HSSP: P23094; INDO.
CC EcoGene: EGL3456; hcae.
CC InterPro: IPR001281; Rieske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00335; Rieske_1.
CC Pfam: PF00848; Ring_hydroxyl_A; 1.
CC PRINTS: PR00090; RINGDIOXYGENASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism: Oxidoreductase; Iron-sulfur; Iron;
CC Dioxygenase; NAD; Complete proteome.
FT METAL 85 85 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 87 87 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 105 105 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

```



```
FT METAL 108 108 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 218 218 IRON (BY SIMILARITY).
FT CONFLICT 20 20 V -> A (IN REF. 1).
FT CONFLICT 384 453 GHARRSKICLEMGLQOQERKRDGIGITNYIFSEIARGM
FT (IN REF. 1).
FT YORWADLSESSEMOEYLDKTAAYQOEVMK -> ATRAPATAN
FT CVMKGLVRRKSAATTAFALTLITISFOKLEPVECTNAGPIE
FT (IN REF. 1).
SQ SEQUENCE 453 AA; 51109 MW; 02535BF5F47643FD CRC64;

Query Match 30.5%; Score 734; DB 1; Length 453;
Best Local Similarity 36.4%; Pred. No. 6,4e-50;
Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGDEELFQHELTIFARNWMLFLTHDLSLPAPGDVYTAKMGIDEIVTSRONDSIRAFNL 79
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 24 IYTDPIYOLELERIGRCMLFLAHESQIPKPDFFNTYMGEDAVVYVVKQKSAIVFLN 83
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 80 VCHHRGKTLVSEAGNAKGFVCSYHGSGFSGNGELOSVPFEKDLYGESLNKKCLGLEKVA 139
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 84 QCHRRAMRVSYADCGNTRAFTCPYHGMVSYGINELIDVLEPRAPYQGLCKSHMGLNEVP 143
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 140 RVSEFGLTYGCFDQAPRLMDYLGDAWYLEPMF-KHSGGLELVGPPGKYVYKAMWKP 198
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 144 CVESYGLIFGMNDTSAPGLRDLIDWYLDGMDLRREGTEIVGVQKWTINCMWKEP 203
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 199 AENFVGDAYVGVTHASSTL-----RSGESIFSSLAGNALP----- 234
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 204 AEGFASDQYHALEFSAVQVYLGAKDQSGDKRLGD-----GQTRAPWETAKDALQFG 256
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 235 --PEGAGLQMTSKYSGGMVLMDCYSGVSHADLVPELMAFGAKOERLKEIGDVARITY 292
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 257 QDHGSGFEFTPEPDANVWV--DGAVSSYRYETAYE-----AQRLGEVYALRL 303
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 293 RSHLNTFVPPNNSMLCSGVFKWNPIDANTEVWYTAIVKMDPEDLKRLADSVQRG 352
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 304 AGHNN--IFPTLSMLNGITLTRYWHPRPDQVEWMAFCITDRAASDEVAAEENSATRAF 361
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 353 GPAGWESDNDNMETASONGKRYOSRSDLSNLGFGEDVYGDVYPPGVKSAIGETS 412
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 362 GPAGFLEQDSEKWCIEQLKRGHARRNSKLCLEMGLOGEKRRDQIPGIT-NYIFSEIA 420
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 413 YKGFYRAYQAHVSSSWAEFEHASTTWHTLTK 445
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 421 ARGMYORWADLSESSEMOEYLDKTAAYQOEVMK 453
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 5
BEDL_PSEPU STANDARD; PRT; 450 AA.
ID BEDL_PSEPU 007944;
AC 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3).
GN BEDCL
OS Pseudomonas putida.
OC Plasmid pHMT112.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC STRAIN=ML2;
RX MEDLINE=93345820; PubMed=8344526;
RA Tan H.-M., Tang H.-Y., Joannou C., Abdel-Wahab N.H., Mason J.R.;
RT "The Pseudomonas putida ML2 plasmid-encoded genes for benzene
dioxygenase are unusual in codon usage and low in G+C content.";
RL Gene 130:33-39(1993).
CC -I- CATALYTIC ACTIVITY: Benzene + NADH + O(2) = cis-1,2-
dihydrobenzene-1,2-diol + NAD(+).
CC -I- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -I- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.
```

```
CC -I- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BEDC1 AND BEDC2), A
CC FERREDOXIN (BEDB) AND A FERREDOXIN REDUCTASE (BEDA).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF148496; AAA1758.1; -.
CC HSSP: P23094; INDO.
CC DR InterPro: IPR001281; Rieske.
CC DR InterPro: IPR001663; Ring_hydroxyl_A.
CC DR Pfam: PF00355; Rieske; 1.
CC DR Pfam: PF00848; Ring_hydroxyl_A; 1.
CC DR PRINTS: PR00090; RINGDIOXYGENASE.
CC DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC KM Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxigenase; NAD; Plasmid.
CC FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC FT METAL 222 222 IRON (BY SIMILARITY).
CC FT METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 51108 MW; 1E0D5E6AC4CF72C8 CRC64;

Query Match 37.0%; Score 649.5; DB 1; Length 450;
Best Local Similarity 34.6%; Pred. No. 2.5e-43;
Matches 148; Conservative 69; Mismatches 174; Indels 37; Gaps 11;
```

```
QY 20 IHGDEELFQHELTIFARNWMLFLTHDLSLPAPGDVYTAKMGIDEIVTSRONDSIRAFNL 79
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 35 IYTDDELYOLELERVAFRSMLLGHETHIRKQGDYFTTYMGEDPVVYVVKQKSAIVFLN 94
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 80 VCHHRGKTLVSEAGNAKGFVCSYHGSGFSGNGELOSVPFEKDLYGESLNKKCLGLEK-- 137
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 95 QCHRRMRICRSDAGNAKAFCTSYHGMAYDTAGNLINVEYEAESFA-----CLDKKEMS 148
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 138 --VARSEFHFGLTYGCFDQAPRLMDYLGDAWYLEPMF-KHSGGLELVGPPGKYVYKAN 194
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 149 PLKARVETTKGLIFANWMDNALIDLTLYGEAKFYMDHMLDKRTBAGTEVYIPGIQKWYIPCN 208
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 195 WKAPAEFVGDAYVHGWMT-HASSLRSGESIFSSLAGNALPPEGAGLQMTSKYSGMGVL 253
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 209 WKFAADQFCSDMYHAGTTHLSGIIAGLPEDLEADLA--PP-----KFGQYRAS 257
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 254 WDGY-SGVHSAAD-----LVPELMAF--GGAKOERLKEIGDVR--ARIYSHLNCYVF 301
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 258 MCGHSGCFYIGDPNNMLAMMGPKVTSYLTEGPAKKAERLGSIRGKIMLEHM--TVF 315
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 302 PNNSMLTGSGVFKWNPIDANTEVWYTAIVKMDPEDLKRLADSVQGTGPPAGWESD 361
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 316 PTCSEFLPGVNTIRTRHMPKRPNEVEWAFVYADAPDDKEEFKQTLTFSGAGVFEED 375
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 362 DNDNMETASONGKRYOSRSDLSNLGFGEDVYGDVYPPGVKSAIGETSYRGFYRAYQ 421
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 376 DGEWVEIGHILRGHKARSRPENAEKMGQVNDPIYDGRISNNVYSEEARGLYAHML 435
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
QY 422 AHVSSSNW 429
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 436 KMTSPDW 443
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 6
TODL_PSEPU STANDARD; PRT; 450 AA.
ID TODL_PSEPU P13450;
AC
```

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toluene 2,3-dioxygenase alpha subunit (EC 1.14.12.-).
GN TODC1.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC STRAIN=FL;
RX MEDLINE=89359301; PubMed=2670929;
RA Zylstra G.J., Gibson D.T.;
RT "Toluene degradation by Pseudomonas putida FL. Nucleotide sequence of
the todC1C2BAD genes and their expression in Escherichia coli.";
RL J. Biol. Chem. 264:14940-14946(1989).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: Toluene degradation; first step.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (TODC1 AND TODC2), A
FERRDOXIN (TODB) AND A FERRDOXIN REDUCTASE (TODA).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04996; AAA26005.1; -
DR PIR: A36516; A36516.
DR HSSP: P23094; IMDO.
DR Interpro: IPR001281; Rleske.
DR Interpro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rleske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXNAS.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxygenase; NAD.
FT METAL 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 IRON (BY SIMILARITY).
FT METAL 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 50944 MW; 038C80F197F3485D CRC64;

Query Match 26.98; Score 647.5; DB 1; Length 450;
Best Local Similarity 34.48; Pred. No. 3.6e-43;
Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;

```

```

DB 264 GFEYDDPMLMLAIIMPKVTYSWTEGPASAKAERLGSVERGSKLMVEHM--TVFPCSFSL 321
OY 308 TCSGFKVWNPIDANTTWTVTYTAIYEKMPEDLKRRLADSQVRGGAGWESDDNDNME 367
DB 322 PGINTVRTMHRGPEVEWMAFTYVADADPDIDKEERROTLRTFSAGVEEDDDEENWV 381
OY 368 TASGCKKYOSRSDLSNLGFGEDVYGDAVYPSVWKSAGTSGYRGRAYOAHVSSS 427
DB 382 EIQHTLRGHRKARSPPFAEMSMQDVTVDNDPYIPORISNNVSEAAAGLYAHMLRMKTSF 441
OY 428 NW 429
DB 442 DW 443

RESULT 7
BPHA_BURCE
ID BPHA_BURCE STANDARD; PRT; 458 AA.
AC P37333;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
dioxygenase).
GN BPHA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB400;
RX MEDLINE=92234948; PubMed=1569021;
RA Erickson B.D., Mondello F.J.;
RT "Nucleotide sequencing and transcriptional mapping of the genes
encoding biphenyl dioxygenase, a multicomponent
polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain
LB400.";
RL J. Bacteriol. 174:2903-2912(1992).
RN [2]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RC STRAIN=LB400;
RX MEDLINE=96011369; PubMed=7592331;
RA Haddock J.D., Gibson D.T.;
RT "Purification and characterization of the oxygenase component of
biphenyl 2,3-dioxygenase from Pseudomonas sp. strain LB400.";
RL J. Bacteriol. 177:5834-5839(1995).
RN [3]
RP ERRATUM.
RA Haddock J.D., Gibson D.T.;
RL J. Bacteriol. 178:2158-2158(1996).
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit.
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
first step.
CC -1- SUBUNIT: Heterohexamer consisting of three bpha subunits and three
bphn subunits. A ferredoxin (bphF) and a ferredoxin reductase
(bphG) must be present to obtain activity.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86348; AAB63425.1; -
DR PIR: B41858; B41858.

```

```
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxxygenase; NAD.
FT INIT_MET 0
FT METAL 99 99 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 122 122 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 232 232 IRON (BY SIMILARITY).
FT METAL 238 238 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51382 MW; AFSEBFD05BAFDF4 CRC64;

Query Match 25.5%; Score 614.5; DB 1; Length 458;
Best Local Similarity 32.9%; Pred. No. 1.4e-40;
Matches 146; Conservative 69; Mismatches 186; Indels 43; Gaps 12;

QY 11 ESLSLQKHLHGDELFQHEKLTIFARNMLFLTHDSLIPAGGYATKAGIDEIVYSRON 70
DB 30 EKGLDPR-IVADQSLYELELELVEFGRSMVLGLGHESHVEPTGDLATYMGEDPVVVRQK 88
QY 71 DGSIRAFNVNCRHRTKTVSEAGNAKGFVCSYHGWGFGSNGELQSVPEKDLGSLNK 130
DB 89 DKSIRKFLNQCRRGRICRSDGNAKAFCTCSHGAYDIAGKLVNVPREKEAFCKKKG 148
QY 131 KCLGLEK-ARVESFHGFTYGCDFQDEAPRLMDYLGDAAWYLEPPEKHS-GGLELV 183
DB 149 DC-GFQKAEWGLQARVATYTKGLVFNAMVDQAPDLETYIGDAPRYMDVMDLRPAQTVAI 207
QY 184 GPRGKVIVANKAPAEENVGDAIHVG-WTHASSLSGSESISSLAGNALLPREG--AGL 240
DB 208 GGNQKVVICNMKEFAEQFCSDMYHAGTTHTLSGLAG-----IPRMDLSQA 255
QY 241 QMTSKYSGKGVLMDSYSGVSHADLPELMAGFGAK-----QERLNKEIGD 286
DB 256 QITRK-GNQFRAMAGHSGWYDERGSLAVGPKVYQYGTGPAELAEQRLGHTGAP 314
QY 287 VRARIYRSLNCTVFPNNSMLTCSGVFKYWNPIDANTTEVTAIYEKMPEDLKRLAD 346
DB 315 VR-RMVGQHM--TIPTSCFLPTFNIRIMHPRGNEIEMVAFITVDADAPAIKEEYRR 371
QY 347 SVQRTGPRAGFMSDDNDNMETASQNGKKYQSDSLSLNGEGEYVGDAYIPGVYGS 406
DB 372 HNIIRNFSAGGVFEQDDGEMWVEIQKLGKAKSQPLNMQMIGRSQGTGHPDPGNYG-Y 430
QY 407 AIGETSYRGFYRAYQAHVSSNMA 430
DB 431 VVAEEARAGYHHMMRMSEPSWA 454

RESULT 8
BPHA.COMTE STANDARD: PRT: 457 AA.
AC 046372:
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Biphenyl dioxxygenase alpha subunit (Ec 1.14.12.18) (Biphenyl 2,3-
DE dioxxygenase).
GN BPHA.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID:285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B-356;
RA MEDLINE-97045812; PubMed-8890734;
RA Sylvestre M., Stiois M., Hurtubise Y., Bergeron J., Ahmad D.,
```

```
RA Shareck F., Barriault D., Guillemette I., Juteau J.M.;
RT "Sequencing of Comamonas testosteroni strain B-356-
RT biphenyl/chlorobiphenyl dioxxygenase genes: evolutionary relationships
RT among Gram-negative bacterial biphenyl dioxxygenases."
RL Gene 174:195-202(1996).
CC -I- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -I- COPACITOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (by
CC similarity).
CC -I- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -I- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase
CC (Bphg) must be present to obtain activity (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGNASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U47637; AAC44526.1; -.
CC DR HSSP: P23094; INDO.
CC DR InterPro: IPR001281; Rieske.
CC DR InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske; 1.
CC DR Pfam: PF00848; Ring_hydroxyl_A; 1.
CC DR PRINTS: PR00090; RINGDIOXGNASE.
CC DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC KM Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxxygenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 457 AA; 51691 MW; D13FC0635FACBF5 CRC64;

Query Match 25.5%; Score 614; DB 1; Length 457;
Best Local Similarity 33.1%; Pred. No. 1.5e-40;
Matches 144; Conservative 73; Mismatches 172; Indels 46; Gaps 13;

QY 20 IHGDELFQHEKLTIFARNMLFLTHDSLIPADGVYATKAGIDEIVYSRONGSIKRAFLN 79
DB 39 IYADODLYOLELELVEFGRSMVLMLGHETHIPKIGDYLTLYMGEDPVIMVQKQSIKVFILN 98
QY 80 VQRHGRKTVSYEAGNAKGFVCSYHGWGFGSNGELQSVPEFEDLYGSEJUNKKCLGLEK- 138
DB 99 QCRHGRMIRIVRSDGNAKAFCTCSHGMAVDIAGNLVNVPEKEAFCKRKGDC-GRDKAD 157
QY 139 -----ARVESFHGFTYGCDFQDEAPRLMDYLGDAAWYLEPPEKHS-KHSGGLVGPGRKVTIK 192
DB 158 WGPLQARVETYGIVFANNDPEAPDLKTYLSDAMPYMDVMDLRTENGTEGTAIGIQKWVTP 217
QY 193 ANWKAPAEENVGDAIHVG-WTHASSLSRSESISSLAGNAALPPEGAGLQMT-----SKYQ 247
DB 218 CNMKFAEQFCSDMYHAGTMSHLSGLVAG-----LPPE---MDLQIQLSKNG 262
QY 248 SGMGVIMDXY-----SGVSHADLPELMAGF--GGAQOEHLNKEIGYR-ARIYRSL 296
DB 263 NQFRSAMGSHGAGWFTINDSSILSVVGPRIYQWVGPAEAKARVPPOLPLIDMEGQHM 322
QY 297 NCTVFPNNSMLTCSGVFKYWNPIDANTTEVTAIYEKMPEDLKRLADSVQRTGPGAG 356
DB 323 --TVPTSCFLGINTIRKWHPRGNEIEMVAFVLVDADAPEDIKEEFLQWIRFPNMG 380
QY 357 EWESDDNDNMETASQNGKKYQSRDLSNLGFGEDVYGDAYIPGVYGSKA--ISETSYR 414
:| | | : : : : : | | | | |
```



DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske\_1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIAGONASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
KM Dioxigenase; NAD.  
FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 222 222 IRON (BY SIMILARITY).  
FT METAL 228 228 IRON (BY SIMILARITY).  
SQ SEQUENCE 448 AA; 50208 MW; 6D1DBD5B2F040BE7 CRC64;  
Query Match 24.3%; Score 586.5; DB 1; Length 448;  
Best Local Similarity 32.9%; Pred. No. 2.1e-38;  
Matches 142; Conservative 68; Mismatches 176; Indels 45; Gaps 11;  
DR 20 IHGDELLFQHELTIPARWMLFLTHDSLIAPADGYTAKMGIDEVYVSQNDGSIATFLN 79  
DB 35 IYDEDLQLELEKVRKRSWLLGHETQIRKPEDYITTYGDDPVVVRKDKASIAVFLN 94  
QY 80 VCRHKGTLVSEAGNAKGFVCSYHMGFGSNGELQSVPEKDLVGEISLNKKGLKEV- 137  
DB 95 QCRHGRMIRCRDAGNAKAFCTYHGAVYTAGNLVNVPEAESFA-----CLNKKEWS 148  
QY 138 --VARSEFHGFIYGCDEAPPLMDYLGDAAWYLEPMF-KHSGGLELVGPGRVVIKAN 194  
DB 149 PLKARVETVYKGLIFANWDENAVLDLTLYLGEAKFYMDHMLDRTEAGTEALRIGVQKWYIPCN 208  
QY 195 WKPAENFVGDAYHVGWT-HASSLSRG--SIFSSLAGNALPREGAGLOMTSKYSGMG 231  
DB 209 WKRAAEFCSDMTYHAGTSHLSGLAGLPDL-----EMADLPVYGVKQYRASKGSGHS 263  
QY 252 VIMDGYSVGHADLVPELMAF--GGAKEOELNKEIGDVR--ARIYSHLNCYTFPNNSML 307  
DB 264 GFYVDGNLMLALMGPRVTSYWGEPASEKAARLSVEGSKLMEWHM--TYFPICPSFL 321  
QY 308 TCGSVKRVNMPIDANTTEVYTAIVEKDMPEDKRRLADSVQ-----RTGGPAGFW 358  
DB 322 PGINTVATLASARAEKEVWAFVVDADADDIKEFERARLRFESPVACSSRTTGRGTG- 379  
QY 359 ESDNDNMENASONGKKYQSRDLSNLGFGHDYGDAYPGVYKSAIGESYGFPR 418  
DB 380 -----SRSTSCSATSRSPFNEMSMQTVNDPVPGRISNNVYSEEAARGLYA 430  
QY 419 AYQAHVSSSNW 429  
DB 431 HMLRMNTSPDW 441  
RESULT 11  
ID BPAL\_PSES1 STANDARD; PRT; 458 AA.  
AC 052438;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Biphenyl dioxigenase alpha subunit (Ec 1.14.12.18) (Biphenyl 2,3-  
dioxigenase).  
GN BPAL.  
OS pseudomonas sp. (strain KKS102).  
OC Bacteria; Proteobacteria.  
OX NCBI\_TaxID=307;  
RN SEQUENCE FROM N.A.  
RP MEDLINE=94324977; PubMed=8048958;  
RA Fukuda M., Yasukouchi Y., Kikuchi Y., Nagata Y., Kimbara K.,  
RA Horichi H., Takagi M., Yano K.;  
RT Identification of the bpha and bphb genes of Pseudomonas sp. strains  
KKS102 involved in degradation of biphenyl and polychlorinated

RT biphenyls".  
RL Biochem. Biophys. Res. Commun. 202:850-856(1994).  
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-  
phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).  
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By  
similarity).  
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;  
first step.  
CC -1- SUBUNIT: Heterohexamer consisting of 3 bpha subunits and 3  
bphb4 subunits. A ferredoxin (bphb3) and a ferredoxin reductase  
(bphb4) must be present to obtain activity (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
DIOXYGENASE ALPHA SUBUNIT FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: D17319; BAA04137.1; .  
CC HSSP: P23094; INDO.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske\_1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIAGONASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
KM Dioxigenase; NAD.  
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 103 103 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 121 121 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 124 124 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 234 234 IRON (BY SIMILARITY).  
FT METAL 240 240 IRON (BY SIMILARITY).  
SQ SEQUENCE 458 AA; 51876 MW; 128B8C38E2A3CFMA CRC64;  
Query Match 24.2%; Score 584; DB 1; Length 458;  
Best Local Similarity 33.6%; Pred. No. 3.4e-38;  
Matches 147; Conservative 65; Mismatches 174; Indels 52; Gaps 14;  
DR 20 IHGDELLFQHELTIPARWMLFLTHDSLIAPADGYTAKMGIDEVYVSQNDGSIATFLN 79  
DB 40 IYADQLEYELERIFARSWLLGHFAHLPKTDGYLTYYGDDPVVVRKDKASIAVFLN 99  
QY 80 VCRHKGTLVSEAGNAKGFVCSYHMGFGSNGELQSVPEKDLVGEISLNKKGLKEV- 138  
DB 100 QCRHGRMIRCRDAGNAKAFCTYHGMAIDYDAGNLVNVPEKAEFCDDKEGDC-GRDKAD 158  
QY 139 -----ARVESFHGFIYGCDEAPPLMDYLGDAAWYLEPMF-KHSGGLELVGPGRVVIK 192  
DB 159 WGPLARVETVYKGLIFANWDENAVLDLTLYLGEAKFYMDHMLDRTEAGTEALRIGVQKWYIPCN 208  
QY 193 AMKKAENFVGDAYHVG-WTASSLSRSESTFSSLAGNALPREGAGLOMT-----SKTG 247  
DB 219 CMMKFAAEQFCSDMYHAGTMAHLISGLV-----SLPPE---MDLTOVQVSKMG 263  
QY 248 SGMGVLMDSYGVHSDLVPELMAFGAK-----QERLNKEIGVRRAR-IYRSLU 296  
DB 264 SQFRAMGSHGSGWFLINDAAILMAVGPITTYWOGPAEKAARLNOMPTOTMFGOH 323  
QY 297 NCTVFPNNSMLTCSGVKYVNPIDANTTEV---WTYAIYEKDMPEDKRRLADSVQRTGG 353  
DB 324 --TVFPTCSFLPGINTIRSMHPRGPNVEECGSPSSMPMR---PEDIKFEFRONIRFN 378  
QY 354 PAGFWESDNDNMENASONGKKYQSRDLSNLGFGHDYGDAYPGVYKSA--IGET 411  
DB 379 AGCTFEQDDGEMWVETIORLGRHAKASAPLCAOMGLNVPNKSNPDFP---GKTAVYAAEE 435  
QY 412 SYRGFYRAYQAHVSSSNW 429

Db 436 AARGYHHWARMSEPNW 453

## RESULT 12

XYLX\_PSEPU ID XYLX\_PSEPU STANDARD; PRT: 454 AA.

AC P23099;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Toluene 1,2-dioxygenase alpha subunit (EC 1.14.12.-).  
GN XYLX.  
OS Pseudomonas putida.  
OC Plasmid TOL PMMO.  
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas  
OX NCBI\_Taxid=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92155191; Pubmed=1740120;  
RA Neidle E.L., Hartnett C., Ornston L.N., Bairoch A., Reik M.,  
RA Harayama S.,  
RT "cis-diol dehydrogenases encoded by the TOL PMMO plasmid xylX gene  
and the acinetobacter calcoaceticus chromosomal band gene are members  
of the short-chain alcohol dehydrogenase superfamily.";  
RL Eur. J. Biochem. 204:113-120(1992).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: DEGRADATION OF BENZOATE AND TOLUATE TO CATECHOL.  
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS, THE  
TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (XYLX AND XYLX'), AND AN  
ELECTRON TRANSFER COMPONENT (XYLZ).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----

DR EMBL: M64747; AAA26047.1; -.  
DR PIR: S23482; S23482.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIOXGNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
KW Dioxxygenase; NAD; Plasmid.  
FT METAL 92 92 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 94 94 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 221 221 IRON (BY SIMILARITY).  
FT METAL 226 226 IRON (BY SIMILARITY).  
SQ SEQUENCE 454 AA; 51898 MW; F68D1C41E236C077 CMC64;

Query Match 17.08; Score 410.5; DB 1; Length 454;  
Best Local Similarity 31.28; Pred. No. 1.2e-24;  
Matches 105; Conservative 64; Mismatches 135; Indels 33; Gaps 14;

QY 23 DEELFOHELKTIFARNMLFLTHDSLIPARGDYVAKMKGIDEIVYSKQNDGSRALANCR 82  
DB 34 DPLRFLEDMHIFEGNWIYLAHESQIPERKNDYTTQKROPFIIRNKDGELENAVNACS 93  
QY 83 HRGKTLVSEAGNAKGFCSYHGMGFGSNGELOSVPFEKDL-YGESLKKKCG--LKEV 138  
DB 94 HRGKTLVSEAGNAKGFCSYHGMGFGSNGELOSVPFEKDL-YGESLKKKCG--LKEV 151

QY 139 ARVESFHGFIYGCDFQDQAPPLMDYLGDAAWYLEPMFKHS-GGLELVGPPEKVIKANKWA 197  
DB 152 ARFASYRGFLFQSLREDVAPLEEFLEGSRKVIDMVDQSPGELVLGSSSYVEGWMKV 211  
QY 198 PAENFVGADHYVG---WTHASS-----LR-SGESIFSSLGNAALPREGAGLQMTSKYGS 248  
DB 212 QVEN-GAGGYHSTVHWVWYAAATQOQRKLRDAGDIDIRAWTA--SSMGDGGGFY--SFEN 265  
QY 249 GNGVLMDCYSGVSHADYLPBELMAFGAGAKQERLNEIGDVRAR-TYRSHLQTVFPPNNSML 307  
DB 266 GHQWYMAKMGDKPNRPLF-----AERDRLASEFGEARADMMIGVSRNLCLTPNLYLM 317  
QY 308 TCSCG-VFKVWNPIDANTTEVWTFYALVEK-DMPEDLKR 342  
DB 318 DOFGSQLRTRPLRPLSVDRREITRYCIAPKGFTPRRARR 354

## RESULT 13

BENA\_ACICA ID BENA\_ACICA STANDARD; PRT: 461 AA.

AC P07769;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10).  
GN BENA.  
OS Acinetobacter calcoaceticus.  
OC Bacteria: Proteobacteria; gamma subdivision; Moraxellaceae;  
OC Acinetobacter.  
OX NCBI\_Taxid=471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=BD413 / ADP1;  
RX MEDLINE=91358314; Pubmed=1885518;  
RA Neidle E.L., Hartnett C., Ornston L.N., Bairoch A., Reik M.,  
RA Harayama S.,  
RT "Nucleotide sequences of the Acinetobacter calcoaceticus benABC genes  
for benzoate 1,2-dioxygenase reveal evolutionary relationships among  
multicomponent oxygenases.";  
RL J. Bacteriol. 173:5385-5395(1991).  
RN [2]

RP REVISIONS TO 84; 103-104; 171-172 AND 380-382.  
RA Eddy D.M., Neidle E.L.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: DEGRADATION OF BENZOATE TO 2-HYDRO-1,2-DIHYDROXYBENZOATE  
(DHB).  
CC -1- CATALYTIC ACTIVITY: Benzoate + NADH + O(2) -> catechol + CO(2) +  
NAD(+).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: DEGRADATION PATHWAY LEADING FROM BENZOATE TO CATECHOL.  
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS: THE  
TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BENA AND BENA'), AND  
AN ELECTRON TRANSFER COMPONENT (BENC).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----

DR EMBL: AF009224; AAC46436.2; -.  
DR PIR: S23477; S23477.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIOXGNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;

Query Match	Best Local Similarity	Matches 128;	Conservative 78;	Mismatch 204;	Indels 52;	Gaps 20;
Query Match	16.4%;	Score 396;	DB 1;	Length 461;		
Best Local Similarity	27.7%;	Pred. No. 1.7e-23;				
Matches 128;	Conservative 78;	Mismatch 204;	Indels 52;	Gaps 20;		
5	NTIIV--SESGLSOAH-LIHGDEELFOHELKTIFAPRNMLFLTHDSLIPAPGYAAKKGI	61				
16	DELYADNTEGTEGFKLHRSVFTDQALFDLEMKITFEGNMWYIALHESQIPNNNDYTTTGTGR	75				
62	DEIVSRONDGSIRAFILVNCVRRHGTIVSVEAGNAKGFVCSYHGMFGSGNGLQSVPEEK	121				
76	QPIILARRNNGELNMINAMINACSHRGAQLCRHKRKNKNTTTCPRHGMVFNNSGKLKLVKDP	135				
122	DL-YEESLNKK-CLGLKEVAVESHGHTIGCFQDEAPPLMDYLGDAAMYLEPMFKHSG-	178				
136	DAGYSDCCFNQSGHDLKVYARESYKGLFQSLNDVSLQEFLETKRIIDIMYQSDQ	195				
179	GLEIVGPGKVKVIRANKMKAPENFEFGDAVYHG---WTHASSL-----RSGSRPSSLAG	229				
196	GLEVLARGVSTYTEBNKWLTAEN-GADGYHSAVMNNAATLQHKKEQAGDTIAMSAG	254				
230	NAALPEBAGLOMTSKYG--SGMGVLMWDGYSGVHSADLVPELM---AFGAKOERLKE	283				
235	SWG--KHGGG-----SYGEFHGMHMLMTQMGPNEDRPNFPKAAEYTEFGAMSKWM--	304				
284	IGDVARIRYRSLNCTVPPNNSMLTCSG-VFKVMNPIDANTTEVTVAYVE-KDMPELK	341				
305	-----IERSR-NICLTPNYIMQFESQIRVLRPISVNKTETVITTCIARPEAPEARA	356				
342	RLADSVORTGGPGPGEWESDDNNKETASQNGKKYQSRDSSLNLTGFGEDVYG--DA--	397				
357	RRIR-QYEDDFENASGMATPDDLEFRACQAGACIELEMMCMCR--GSKHMTVGDPDAAN	413				
398	--VYPGVGKSAIGETSYRGFYRAYQAHVSSNMNAEEFHAS	436				
414	EIGLKPAISGIKTEDEGLYLAHQHVLKSKMOAIAAEKERAS	455				
RESULT 14						
YEAM_ECOLI						
ID YEAM_ECOLI	STANDARD:	PRT:	374 AA.			
AC P76253;						
DT 15-JUL-1998 (Rel. 36, Created)						
DT 15-JUL-1998 (Rel. 36, Last sequence update)						
DT 16-OCT-2001 (Rel. 40, Last annotation update)						
DE Putative dioxigenase alpha subunit Yeam (Ec 1.14.1.1.).						
DE YEAM OR B1802 OR Z2845 OR ECS2511.						
OS Escherichia coli, and						
OS Escherichia coli 0157:H7.						
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;						
OC Escherichia.						
NCBI_TaxID=562, 83334;						
OX (1)						
RP SEQUENCE FROM N.A.						
RC STRAIN=K12 / MG1655.						
RX MEDLINE=97426617; PubMed=9278503;						
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,						
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,						
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,						
RA Mau B., Shao Y.;						
RT "The complete genome sequence of Escherichia coli K-12.";						
RL Science 277:1453-1474(1997).						
[2]						
RP SEQUENCE FROM N.A.						
RC STRAIN=K12;						

```

RX MEDLINE=97251358; PubMed=9097040;
RA Itch T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nasimubek H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasundaram S., Tagami T., Takeda J., Takemoto K., Mada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40,1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
RN [3]
RP
RC SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hayekett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoustis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [4]
RP
RP SEQUENCE FROM N.A.
RX STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11256796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shimagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001)
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- SUBUNIT: PROBABLE HETERODIMER OF YEAW AND YEAW.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC CC THE SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000274: AAC74872.1; -
DR EMBL: D90823: BAA15597.1; ALT: INT.
DR EMBL: D90824: BAA15606.1; ALT: INT.
DR EMBL: AE005403: AAG56791.1; -.
DR EMBL: AP002558: BAB35934.1; -.
DR EcoGene: EGI3509; yeaw.
DR InterPro: IPR001281; Riese.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; FALSE NEG.
KW Hypothetical protein; Oxidoreductase; Iron-sulfur; Iron; Dioxygenase;
FT METAL 89 89 IRON-SULFUR (2FE-2S) (BY SIMILARITY) .
FT METAL 91 91 IRON-SULFUR (2FE-2S) (BY SIMILARITY) .
FT METAL 109 109 IRON-SULFUR (2FE-2S) (BY SIMILARITY) .
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY) .
FT METAL 211 211 IRON (BY SIMILARITY) .
FT METAL 216 216 IRON (BY SIMILARITY) .
SQ SEQUENCE 374 AA; 42561 MW; BB5386ACA9585606 CRC64;
Query Match 10.1%; Score 244; DB 1; Length 374;
Best Local Similarity 26.9%; Pred. No. 9.3e-11;
Matches 97; Conservative 52; Mismatches 122; Indels 90; Gaps 19;

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:55:37 ; Search time 13.1667 Seconds  
(without alignments)  
3278.305 Million cell updates/sec

Title: US-09-843-250-32

Perfect score: 2410  
Sequence: 1 MNVNNKILVSESGLSQKHLLI.....AEPEHASTWHELTFTKTR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: pirl:\*\*\*  
2: pirl:\*\*\*  
3: pirl:\*\*\*  
4: pirl:\*\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2401	99.6	449	2 JS0071	naphthalene dioxyg
2	2401	99.6	462	2 S27632	naphthalene 1,2-di
3	2343	97.2	449	2 JN0644	naphthalene 1,2-di
4	2289	95.0	449	2 C53217	polycyclic aromati
5	2048	85.0	447	2 JC5352	2-nitrotoluene dio
6	954	39.6	459	2 T31134	naphthalene dioxyg
7	771	32.0	450	2 T31256	terminal oxygenase
8	734	30.5	453	2 A65031	biphenyl dioxygena
9	734	30.5	453	2 D91054	biphenyl dioxygena
10	734	30.5	453	2 G85898	biphenyl dioxygena
11	722.5	30.0	455	2 T31258	biphenyl dioxygena
12	649.5	27.0	450	1 JN0812	aromatic oxygenase
13	647.5	26.9	450	1 A36516	benzene 1,2-dioxyg
14	636	26.4	461	2 S51757	toluene dioxygens
15	614.5	25.5	459	1 B41858	biphenyl dioxygena
16	614	25.5	457	1 JCA4993	biphenyl dioxygena
17	614	25.5	458	1 A42409	biphenyl dioxygena
18	602	25.0	431	2 JN0098	carbazole dioxygen
19	586.5	24.3	448	1 A29830	benzene 1,2-dioxyg
20	584	24.2	458	2 JCA2647	biphenyl dioxygena
21	449	18.6	469	2 T50934	dioxygenase D1A1,
22	415	17.2	471	2 T31281	benzoate 1,2-dioxy
23	410.5	17.0	454	2 A41659	benzoate 1,2-dioxy
24	406.5	16.9	455	2 E83332	toluene 1,2-dioxyg
25	391	16.2	461	2 S23477	probable benzoate
26	389.5	16.2	464	2 G83331	anthranilate dioxy
27	333.5	13.8	424	2 E83384	probable ring-hydr
28	327	13.6	426	2 T31278	biphenyl dioxygena
29	311.5	12.9	468	2 G97447	hypothetical prote

30	311.5	12.9	468	2 AH2665	ring hydroxylating
31	309.5	12.8	391	2 T31251	aromatic oxygenase
32	291	12.1	429	2 F82970	probable ring hydr
33	287	11.9	412	2 H95311	probable aromatic
34	285.5	11.8	415	2 AG3320	benzoate 1,2-dioxy
35	269.5	11.2	404	2 H87635	Rieske 2Fe-2S faml
36	260	10.8	420	2 T31285	biphenyl dioxygena
37	248.5	10.3	374	2 AF0304	probable dioxygena
38	244	10.1	374	2 C85791	probable choline m
39	244	10.1	374	2 G90942	probable choline m
40	244	10.1	374	2 B64941	probable choline m
41	233.5	9.7	382	2 G70946	probable dioxygena
42	224	9.3	446	2 T14542	choline monooxygen
43	223	9.3	439	2 T09214	choline monooxygen
44	214.5	8.9	186	2 JCS354	2-nitrotoluene dio
45	186	7.7	426	2 T08550	choline monooxygen

## ALIGNMENTS

```
RESULT 1
JS0071
naphthalene dioxygenase (Ec 1.14.12.-) ndoB protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Aug-1999
C:Accession: JS0071
R:Kurkela, S.; Lehtvaestalo, H.; Palva, E.T.; Teeri, T.H.
Gene 73, 355-362, 1988
A:Title: Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase and oxidoreductase; Rieske iron-sulfur protein
A:Reference number: JS0070; MUID:89211973; PMID:3243438
A:Accession: JS0071
A:Molecule type: DNA
A:Residues: 1-449 <KUR>
A:Cross-References: GB:M23914; NID:9151392; PIDN:AAB47591.1; PID:9151394
C:Comment: Naphthalene dioxygenase system is composed of three proteins.
C:Genetics:
A:Gene: ndoB
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match          99.6%  Score 2401:  DB 2:  Length 449:
Best Local Similarity 99.8%  Pred. No. 4,8e-183;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	1	MNVNNKILVSESGLSQKHLLIHGDEELFQHELTIFARNMLFTLHDSLTLPAPDYVTAKMG	60
DB	1	MNVNNKILVSESGLSQKHLLIHGDEELFQHELTIFARNMLFTLHDSLTLPAPDYVTAKMG	60
QY	61	IDEVIVSRNDGSIIRAFVNCRRGKTLVSVAGNKKGVCSYHGMGRSGNCELOSVPPE	120
DB	61	IDEVIVSRNDGSIIRAFVNCRRGKTLVSVAGNKKGVCSYHGMGRSGNCELOSVPPE	120
QY	121	KDLVGSLSNKKCLGKEVARVSEFGEITGCDQEPPLMDVLGDAWYLEPMFKHSGGL	180
DB	121	KDLVGSLSNKKCLGKEVARVSEFGEITGCDQEPPLMDVLGDAWYLEPMFKHSGGL	180
QY	181	ELVGPGRKVIVIRANKKAPAEVFDAYHVHGWTHASSLSRSGESIFSSLAGNMAALPREGAGL	240
DB	181	ELVGPGRKVIVIRANKKAPAEVFDAYHVHGWTHASSLSRSGESIFSSLAGNMAALPREGAGL	240
QY	241	QMTSKYSGMGVLMGYSVSHADLVPELMAFGAKKOEKRLNKETDVARIRYSLNCTV	300
DB	241	QMTSKYSGMGVLMGYSVSHADLVPELMAFGAKKOEKRLNKETDVARIRYSLNCTV	300
QY	301	FPNNSMLTCSGVFKVWNPIDANTTEVMVYATIEKMPDILKRLADSVORGGPAGFES	360
DB	301	FPNNSMLTCSGVFKVWNPIDANTTEVMVYATIEKMPDILKRLADSVORGGPAGFES	360
QY	361	DDNDMMETASONGKRYQSRSDLSNLGFGEDVYGDAYPGVNGSAIGETSYRGVFRAY	420
DB	361	DDNDMMETASONGKRYQSRSDLSNLGFGEDVYGDAYPGVNGSAIGETSYRGVFRAY	420

Db 361 DDNDMETSANGKKYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 420

Qy 421 QAHVSSNMMAEFERHASTWHTLTKTTDR 449

Db 421 QAHVSSNMMAEFERHASTWHTLTKTTDR 449

## RESULT 2

S27632  
naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large chain  
C:Species: Pseudomonas sp.  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S27632; B49343  
R:Denome, S.A.; Young, K.D.  
Submitted to the EMBL data library, February 1992  
A:Description: Cloning and molecular characterization of genes involved in metabolism of  
A:Reference number: S27631  
A:Accession: S27632  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <DEN>  
A:Cross-references: EMBL:M60405  
R:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
J. Bacteriol. 175, 6890-6901, 1993  
A:Title: Metabolism of dibenzothioophene and naphthalene in Pseudomonas strains: complete  
A:Reference number: A49343; MUID:94042852; PMID:822631  
A:Accession: B49343  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 14-462 <DE2>  
A:Cross-references: GB:M60405; NID:9151195; PIDN:AAA16125.1; PID:9294351  
A:Experimental source: strain C18  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:84-132/Domain: Rieske [2Fe-2S] homology <RSK>  
F:94,96,114,117/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 99.6%; Score 2401; DB 2; Length 462;  
Best Local Similarity 99.8%; Pred. No. 5e-183;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MNYNKKILVSSGSLQKHLHGDEBLFOHELKTIFARNWLEFLTHDSLIPAPGDVYTKMG 60  
Db 14 MNYNKKILVSSGSLQKHLHGDEBLFOHELKTIFARNWLEFLTHDSLIPAPGDVYTKMG 73  
Qy 61 IDEVIVSRONGSTRAFLNVCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
Db 74 IDEVIVSRONGSTRAFLNVCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 133  
Qy 121 KDLYGESLNKKCLGLEKVARVESFHGFTYGCFODEAPPLMDYLGDAWYLEPMFKHSGGL 180  
Db 134 KDLYGESLNKKCLGLEKVARVESFHGFTYGCFODEAPPLMDYLGDAWYLEPMFKHSGGL 193  
Qy 181 ELVPPGKVVYIKAMWKAENFVGDAYHVGTTHASSLSRSGESIFSSLGNAALPREGAGL 240  
Db 194 ELVPPGKVVYIKAMWKAENFVGDAYHVGTTHASSLSRSGESIFSSLGNAALPREGAGL 253  
Qy 241 QMTSKYSGMGVLDGYSVGSADLVPFLMAFGAKORLNKEIGDVARITYSHLNCV 300  
Db 254 QMTSKYSGMGVLDGYSVGSADLVPFLMAFGAKORLNKEIGDVARITYSHLNCV 313  
Qy 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYALVEKDMPEDLKRLADSVORTGPAFWES 360  
Db 314 FPNNSMLTCSGVFKVWNPIDANTTEVWTYALVEKDMPEDLKRLADSVORTGPAFWES 373  
Qy 361 DDNDMETSANGKKYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 374 DDNDMETSANGKKYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 433  
Qy 421 QAHVSSNMMAEFERHASTWHTLTKTTDR 449  
Db 434 QAHVSSNMMAEFERHASTWHTLTKTTDR 462

## RESULT 3

JN0644  
naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large ch  
N:Alternate names: nahc protein  
C:Species: Pseudomonas putida  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
C:Accession: JN0644  
R:Simon, M.J.; Osslund, T.D.; Saunders, R.; Enstley, B.D.; Suggs, S.; Harcourt, A.; Su  
Gene 127, 31-37, 1993  
A:Title: Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida st  
A:Reference number: JN0640; MUID:93252277; PMID:8486285  
A:Accession: JN0644  
A:Molecule type: DNA  
A:Residues: 1-449 <STM>  
A:Cross-references: GB:M83949; NID:9151384; PIDN:AAA25902.1; PID:9151387  
C:Comment: This protein is a member of naphthalene dioxygenase multicomponent enzyme  
C:Genetics:  
A:Gene: nahc  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2  
C:Keywords: 2Fe-2S; heterotrimer; metalloprotein; oxidoreductase; Rieske iron-sulfu  
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 97.2%; Score 2343; DB 2; Length 449;  
Best Local Similarity 96.7%; Pred. No. 2e-178;  
Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MNYNKKILVSSGSLQKHLHGDEBLFOHELKTIFARNWLEFLTHDSLIPAPGDVYTKMG 60  
Db 1 MNYNKKILVSSGSLQKHLHGDEBLFOHELKTIFARNWLEFLTHDSLIPAPGDVYTKMG 60  
Qy 61 IDEVIVSRONGSTRAFLNVCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
Db 61 IDEVIVSRONGSTRAFLNVCRRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
Qy 121 KDLYGESLNKKCLGLEKVARVESFHGFTYGCFODEAPPLMDYLGDAWYLEPMFKHSGGL 180  
Db 121 KELYGESLNKKCLGLEKVARVESFHGFTYGCFODEAPPLMDYLGDAWYLEPMFKHSGGL 180  
Qy 181 ELVPPGKVVYIKAMWKAENFVGDAYHVGTTHASSLSRSGESIFSSLGNAALPREGAGL 240  
Db 181 ELVPPGKVVYIKAMWKAENFVGDAYHVGTTHASSLSRSGESIFSSLGNAALPREGAGL 240  
Qy 241 QMTSKYSGMGVLDGYSVGSADLVPFLMAFGAKORLNKEIGDVARITYSHLNCV 300  
Db 241 QMTSKYSGMGVLDGYSVGSADLVPFLMAFGSKORLNKEIGDVARITYSHLNCV 300  
Qy 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYALVEKDMPEDLKRLADSVORTGPAFWES 360  
Db 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYALVEKDMPEDLKRLADSVORTGPAFWES 360  
Qy 361 DDNDMETSANGKKYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 361 DDNDMETSANGKKYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Qy 421 QAHVSSNMMAEFERHASTWHTLTKTTDR 449  
Db 421 QAHVSSNMMAEFERHASTWHTLTKTTDR 449

## RESULT 4

C55217  
polycyclic aromatic hydrocarbon dioxygenase (EC 1.14.12.-) iron-sulfur protein large  
C:Species: Pseudomonas putida  
C:Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 20-Jun-2000  
C:Accession: C55217  
R:Takizawa, N.; Kaide, N.; Torioge, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara  
J. Bacteriol. 176, 2444-2449, 1994  
A:Title: Identification and characterization of genes encoding polycyclic aromatic hy  
A:Reference number: A55217; MUID:94209249; PMID:8157615  
A:Accession: C55217  
A:Status: preliminary

Best Local Similarity 83.7%; Pred. NO. 5.7e-155;  
Matches 376; Conservative 34; Mismatches 37; Indels 2; Gaps

187 GKVVIKANKKAPAEENVGDAYHYGHTHASSLRSGESIFFSLAGNNA-LPEEGAGLOMTSK 245

Db 186 LKSTLACNMKVPTEPNVGDYGVGTHAALQMIGELAGISGNRAMPFDLLGLQFTMR 245  
QY 246 YGSGMGLMDYSGVH--SADLPELMAFGAKOERLNKEIGDVARIRYRSHLNCTVPPN 303  
Db 246 HGHGFGILDNAATAIHVKRQGVYKLEETRGCIKEKPE----REKLYGHMNSTILPN 301  
QY 304 NSMLTCSGVFKVMPIDANTTEWMTYAIVEKMPEDLKRRLADSVQRTGAPGEMESDN 363  
Db 302 CSFLYGTNFEKIMWRPHEIEVWMTYVPRKADTEKRSIOREAIRSFGAGTLESDDG 361  
QY 364 DNMETAQNGKKYOSRSDLSNLFGEEDVYG-DAVYGVYKSAIGSTYRGFYRAQA 422  
Db 362 ENMSAYNNNGNITTRCRM--NSSMGKDRGPPHYGVIGVSGFTGTYRGFYRMOE 419  
QY 423 HVSSNMAEFEEHASTW 439  
Db 420 MLDAPMAAIRANDTW 436

## RESULT 7

T31256  
terminal oxygenase component large chain homolog - Sphingomonas aromaticivorans plasmid  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31256  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.;  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A:Reference number: Z20992

A:Accession: T31256  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-450 <ROM>  
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378397; PIDN:AD03980.1  
C:Genetics:

A:Genome: plasmid pML1  
A>Note: DpHAla  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 32.08; Score 771; DB 2; Length 450;

Best Local Similarity 36.18; Pred. No. 2,Je-53;

Matches 157; Conservative 88; Mismatches 184; Indels 6; Gaps 6;

QY 1 MNTNKLIVSESGSLQKHLHGDELFQHELTIFARNMLFLTHDSLIPAPGDVYTAAG 60  
Db 1 MNDIADLVSRIGRQSRSTIASEDIYRQELERIFGRCMFLVHTSOLPKFQDIFRTFMG 60  
QY 61 IDEVIVSRQDGSIRAFILNVCRRHKGTLVSEAGNAKGFVCSYHGWFSGNGLQSVPE 120  
Db 61 EDDVIVIRQDGSIRAFILNVCRRHKGTLVSEAGNAKGFVCSYHGWFSGNGLQSVPE 120  
QY 121 KDLYGESLANKKGLGKEVARYESFHGFLYCGFDQDAPPLMDYLGDAWYLEP-MFKHSGG 179  
Db 121 NEAYIFGELDKTKGLLVTYVAEYKGLVFGCMANSPSLDYLDKACFLDVLWDAMPGG 180  
QY 180 LEIVAPRGKVVIRAKMAPENFVGDAYHYGWTASSLRSEST-FSSLA-GNALPREG 237  
Db 181 SALIGETQKAVNLGTNMLPEVNCGGDGHGMAHAGAAVQSDMLDGLSGNSGVDLDG 240  
QY 238 AGLOMTSKYSGMGVLDGYSVHSAIDLVPBELMAFGAKOERLNKEIGDVARIRYRSHL 296  
Db 241 -GLSVAGMGNHVLASLDGVSAYFPDPKRLILEYLEANRQTVIDRLGEVGRQVWAGV 299  
QY 297 NCIVFPPNNSMLTCSGVFKVMPIDANTTEWMTYAIVEKMPEDLKRRLADSVQRTGAPG 356  
Db 300 NITIFPRLQLLPGLMWRVYHPRKPGQIEQWTMAEMEDPEAKVQQLLEMOCTLFGLAG 359  
QY 357 FWSDDDNKNETASONGKKYOSRSDLSNLFGEEDVYGDAVYGVYKSAIGSTYRGF 416  
Db 360 LFDNDGDNLCTACTEBSRGMTAOMDYTTNMAIGRSKRG-PPGDIAAGLVSEHNRXF 418

QY 417 YRAYQAHVSSNMMAE 431  
Db 419 YRMQEHMAETMAE 433

## RESULT 8

A65031  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escheri  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: A65031; S49292  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Siao, Y.  
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A65031

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-453 <BLAT>

A:Cross-references: GB:AE000340; GB:U00096; NID:g1788883; PIDN:AC75591.1; PID:g17888

A:Experimental source: strain K-12, substrain MG1655

R:Tullin, E.; Gasser, F.; Biville, F.

A:Submitted to the EMBL Data Library, September 1994

A:Description: Cloning and sequencing of an E. coli gene homologous to dioxygenase of

A:Reference number: S49292

A:Accession: S49292

A:Molecule type: DNA

A:Residues: 1-19, 'A', 21-383, 'ATAPATANCWKW', 397-398, 451, 'R', 453, 'SANTATIALTISFOKLPL

A:Cross-references: EMBL:Z37966; NID:g350595; PIDN:CAH86018.1; PID:g350595

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [

C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein

F:75-123/Domain: Rieske [2Fe-2S] homology <RSK>

F:85,87,105,108/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 30.58; Score 734; DB 2; Length 453;

Best Local Similarity 36.43; Pred. No. 1,8e-50;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGDELFQHELTIFARNMLFLTHDSLIPAPGDVYTAAGGIDEVIVSRQDGSIRAFILN 79  
Db 24 IYTPDPIQLERLIFGRCMFLVHTSOLPKFQDIFRTFMG 83  
QY 80 VCRHKGTLVSEAGNAKGFVCSYHGWFSGNGLQSVPEKDYGESLANKKGLGKEVA 139  
Db 84 QCRIRAMKVSADCGNTRAFICPRHGSYGINGLIDVLEPRAPYQGLCKSHMGLEVP 143  
QY 140 RVSEFHGFLYCGFDQDAPPLMDYLGDAWYLEP-MFKHSGGLELVGPPGKVVIRAKMAP 198  
Db 144 CVESYKGLIFGNMPTSAPLRDYLGDIAMYLDDGLDRREGTEIVGVQKVINCMWKEP 203  
QY 199 AENFVGAYHYGWTASSLRSEST-FSSLA-GNALPREG 234  
Db 204 AEQFASQYIHALFHASAVOYLAKKDDGSKRLGD-----GQIAPRWETAKDALQFG 256  
QY 235 --PPGAGLOMTSKYSGMGVLDGYSVHSAIDLVPBELMAFGAKOERLNKEIGDVARIRY 292  
Db 257 QDGHSGGFFEFKEDAVWV--DCAVSSYRYETAE-----ABGRLEGEVALRL 303  
QY 293 RSHLNCIVFPPNNSMLTCSGVFKVMPIDANTTEWMTYAIVEKMPEDLKRRLADSVQRTG 352  
Db 304 AGHNN--IFPTLSWLNCTATLRVHPRGPRQVVEWMAFCIDDKAASDEVKAAEFNSATRAF 361  
QY 353 GPAGFLEDDSDENNCETQKLKIGHRANRSKLCLEMGLGQEKRRDDIGPIT--NYIFSETA 420  
Db 362 GPAGFLEDDSDENNCETQKLKIGHRANRSKLCLEMGLGQEKRRDDIGPIT--NYIFSETA 420  
QY 413 YRGFYRAYQAHVSSNMMAEFEEHASTWHTLTX 445  
Db 421 ARGMYQRMADLSSSEWQVYLDKTAAYQGVMK 453

## RESULT 9



[illegible]

Db 209 WKFAEAFQCSMDYHAGTTHLGSITAGLPEDDELADLA--PP-----KFGQRYAS 257

QY 254 WDOY-SCVHSAD-----LVPELMAF--GGAQERLKEIGDVR--ARIYRSLNCTVF 301

Db 258 WGHGSGGFYIGDPNNMLAMGPKVTSYLTGCPAAEKAERLGSIRGTHIMLEHW--TVF 315

QY 302 PNNSMITGSGVFKVNNPIDANTTEWTVTAIVKDMPEDKRLRLADSVORTGPGAFWESD 361

Db 316 PTCSTLPGVNTRIHTMHPGPNPEVEWAFVTVADAPDDIKEFFRQTLITFSAGGVFQD 375

QY 362 DNDNMTASQNGKKYQSRDSDLNLGFGEDVYGDVAPGVVGSKAISGTSRGFYRAYQ 421

Db 376 DGNWVEIHLIRGHKARSRFPNAEKSMQGTVDNDPIYPRISNNVYSEEARGLYAHML 435

QY 422 AHVSSSNW 429

Db 436 KMWTSFDW 443

Query	Match	Best Local Similarity	Score	DB 1:	Length
Db	35	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	20	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	35	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	80	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	95	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	138	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	149	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	195	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	209	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	252	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	264	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	308	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	322	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Qy	366	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;
Db	382	145	Conservative	71	Mismatches 181; Indels 25; Gaps 9;







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:46:47 : Search time 29.1667 Seconds  
(without alignments)  
2051.298 Million cell updates/sec

Title: US-09-843-250-32  
Perfect score: 2410  
Sequence: 1 MNVNNKILVSESGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2410	100.0	449	21	AA12578
2	2404	99.8	449	21	AA12579
3	2402	99.7	449	21	AA12580
4	2402	99.7	449	21	AA12583
5	2401	99.6	449	21	AA12565
6	2401	99.6	449	21	AA12566
7	2401	99.6	449	21	AA12567
8	2401	99.6	449	21	AA12577
9	2400	99.6	449	21	AA12581
10	2400	99.6	449	21	AA12582

11	2396	99.4	448	21	AA12584
12	2396	99.4	449	21	AA12568
13	2391	99.2	449	21	AA12569
14	2343	97.2	449	21	AA12570
15	2315	96.1	449	21	AA12571
16	2289	95.0	449	21	AA12572
17	2216	92.0	449	21	AA12573
18	2186	90.7	447	21	AA12574
19	2049	85.0	447	21	AA12575
20	1965.5	81.6	451	21	AA12576
21	1092	45.3	452	21	AA12577
22	734	30.5	453	22	AA12578
23	660.5	27.3	443	23	AA12579
24	658.5	27.3	898	21	AA12580
25	620.5	25.7	459	16	AA12581
26	615	25.5	458	21	AA12582
27	612	25.4	458	21	AA12583
28	551.5	22.9	427	14	AA12584
29	377.5	15.7	385	22	AA12585
30	376	15.6	490	22	AA12586
31	357	14.8	497	22	AA12587
32	333.5	13.8	424	22	AA12588
33	239	9.9	435	22	AA12589
34	235	9.8	433	22	AA12590
35	235	9.8	433	22	AA12591
36	233	9.7	438	23	AA12592
37	224	9.3	446	19	AA12593
38	223	9.3	439	19	AA12594
39	196	8.1	99	23	AA12595
40	186	7.7	426	23	AA12596
41	181	7.5	35	19	AA12597
42	181	7.5	1193	22	AA12598
43	173	7.2	35	19	AA12599
44	164	6.8	35	19	AA12600
45	126	5.2	405	19	AA12601

## ALIGNMENTS

RESULT 1	
AA12578	
ID	AA12578 standard; Protein; 449 AA.
XX	
AC	AA12578;
ID	
DT	09-NOV-2000 (first entry)
DE	
XX	
XX	Naphthalene dioxygenase mutant F352G protein sequence SEQ ID NO:32.
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KW	Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
OS	Pseudomonas sp.
OS	Synthetic.
XX	
PN	WO200037480-A1.
XX	
PD	29-JUN-2000.
XX	
PF	26-OCT-1999; 99WO-US25079.
XX	
PR	26-OCT-1998; 98US-0105575.
XX	
PA	(IOWA ) UNIV IOWA RES FOUND.
XX	
PI	Parales R, Gibson D, Resnick S, Lee K;
XX	
DR	WPI; 2000-452174/39.
DR	N-PSDB; AAA65352.
XX	
PT	Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT

Claim 9: Page 126-127; 151pp; English.

XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothioephene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA;

Query Match 100.0%; Score 2410; DB 21; Length 449;  
Best Local Similarity 100.0%; Pred. No. 1,4e-224;  
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYNNKILIVSSGSLQKLIHGDELFQHELTIFARWMLFLTHDSLIPAGDYVYAKMG 60  
DB 1 MNYNNKILIVSSGSLQKLIHGDELFQHELTIFARWMLFLTHDSLIPAGDYVYAKMG 60  
QY 1 IDEVIVSRQNDGSIIRAFILNVCRRHGKTLVSYEAGNAGFCVSYHGWFSGNGELQSVPE 120  
DB 61 IDEVIVSRQNDGSIIRAFILNVCRRHGKTLVSYEAGNAGFCVSYHGWFSGNGELQSVPE 120  
QY 121 KDLGESLNKKCLGKEVARVESFHGFIYGCEDAPPLMDYLGDAAWYLEPMFKHSGGL 180  
DB 121 KDLGESLNKKCLGKEVARVESFHGFIYGCEDAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPPEKVVYIKAMWKAPENFVGDAYHVGTWTHASSLSGSEIFSSLAGNALPREGAGL 240  
DB 181 ELVGPPEKVVYIKAMWKAPENFVGDAYHVGTWTHASSLSGSEIFSSLAGNALPREGAGL 240  
QY 241 QMTSKYSSGMVLMDGYSVGSADLVPFLMAFGAKOERLNKEIDVARIYRSLNCTV 300  
DB 241 QMTSKYSSGMVLMDGYSVGSADLVPFLMAFGAKOERLNKEIDVARIYRSLNCTV 300  
QY 301 PPNNSMLTCSGVFVWMPIDANTTEVWTYAIVEKDMPEDLKRLADSVQRTGPGAFWES 360  
DB 301 PPNNSMLTCSGVFVWMPIDANTTEVWTYAIVEKDMPEDLKRLADSVQRTGPGAFWES 360  
QY 361 DDNDNMETASONGKRRYOSRSDLSNLGFGEDVYGDAYPGVYGSAGGYSRYEFYAY 420  
DB 361 DDNDNMETASONGKRRYOSRSDLSNLGFGEDVYGDAYPGVYGSAGGYSRYEFYAY 420  
QY 421 QAHVSSNMAFEHASSTWHTLTKTTDR 449  
DB 421 QAHVSSNMAFEHASSTWHTLTKTTDR 449

RESULT 2  
AAB12579  
ID AAB12579 standard; Protein: 449 AA.  
XX  
AC AAB12579;

XX 09-NOV-2000 (first entry)  
XX  
XX Naphthalene dioxygenase mutant F352A protein sequence SEQ ID NO:33.  
XX

XX Pseudomonas sp. strain NCIB 9616-4; naphthalene dioxygenase: NDO;  
XX inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
XX polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX

OS Pseudomonas sp.  
OS Synthetic.

PN WO200037480-A1.

PD 29-JUN-2000.

PE 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB; AAA65353.

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT

PS Claim 9: Page 128-129; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothioephene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA;

Query Match 99.8%; Score 2404; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 5.3e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNNKILIVSSGSLQKLIHGDELFQHELTIFARWMLFLTHDSLIPAGDYVYAKMG 60  
DB 1 MNYNNKILIVSSGSLQKLIHGDELFQHELTIFARWMLFLTHDSLIPAGDYVYAKMG 60  
QY 61 IDEVIVSRQNDGSIIRAFILNVCRRHGKTLVSYEAGNAGFCVSYHGWFSGNGELQSVPE 120  
DB 61 IDEVIVSRQNDGSIIRAFILNVCRRHGKTLVSYEAGNAGFCVSYHGWFSGNGELQSVPE 120  
QY 121 KDLGESLNKKCLGKEVARVESFHGFIYGCEDAPPLMDYLGDAAWYLEPMFKHSGGL 180  
DB 121 KDLGESLNKKCLGKEVARVESFHGFIYGCEDAPPLMDYLGDAAWYLEPMFKHSGGL 180

DB 121 KDLYGESLKKCLGKEVARVESFHGFTYGCDFDQAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPBGKVIKANKKAPENFVGDAHYHGWTHASSLSRSGESIFSSLAGNALLPREGAGL 240  
DB 181 ELVGPBGKVIKANKKAPENFVGDAHYHGWTHASSLSRSGESIFSSLAGNALLPREGAGL 240  
QY 241 QMTSKYSGSGMGLMDGYSGVSHSADLVPELMARFGAKQERLNEIGDVRARIYRSHLNCV 300  
DB 241 QMTSKYSGSGMGLMDGYSGVSHSADLVPELMARFGAKQERLNEIGDVRARIYRSHLNCV 300  
QY 301 FPNNSMLTCSGVEFKYKWNPIDANTTEWYTAIVEKMPEDLKRRLADSVQRTGPGAFWES 360  
DB 301 FPNNSMLTCSGVEFKYKWNPIDANTTEWYTAIVEKMPEDLKRRLADSVQRTGPGAFWES 360  
QY 361 DDNDMMETASQNGKKYQSDLSNLGFGEDYGDAYPGVGSALGETSYRGFYRAY 420  
DB 361 DDNDMMETASQNGKKYQSDLSNLGFGEDYGDAYPGVGSALGETSYRGFYRAY 420  
QY 421 QAHVSSNMMAEFHASTWHTELTKTTDR 449  
DB 421 QAHVSSNMMAEFHASTWHTELTKTTDR 449

RESULT 3  
AAB12580  
ID AAB12580 standard; Protein: 449 AA.  
XX  
AC AAB12580;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352T protein sequence SEQ ID NO:34.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
XX  
DR N-PSDB; AAA65354.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
XX Claim 9; Page 129-131; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
XX related complex (I) comprising several polypeptides which contain an  
XX alpha subunit that contains substituted amino acids at specific  
XX positions. The polypeptides and host cells are useful for preparing  
XX (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
XX dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
XX dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
XX involves contacting them with naphthalene, biphenyl, phenanthrene,  
XX indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
XX the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
XX dihydriophenanthrene or 3,4-dihydroxy-3,4-dihydriophenanthrene which

CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene, 1,4-dioxan,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzofluorene, 9,10-dihydroanthracene, or 9,10-dihydriophenanthrene  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX  
SQ Sequence 449 AA:  
QY  
DB 1 MNYNKKIIVSESGLSQKHLIHODEELFQHELKTI FARMMFLTHDSLIPAPDVTAKMG 60  
DB 1 MNYNKKIIVSESGLSQKHLIHODEELFQHELKTI FARMMFLTHDSLIPAPDVTAKMG 60  
QY 61 IDEVIVSRQNDGSIRAFINVCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
DB 61 IDEVIVSRQNDGSIRAFINVCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
QY 121 KDLYGESLKKCLGKEVARVESFHGFTYGCDFDQAPPLMDYLGDAAWYLEPMFKHSGGL 180  
DB 121 KDLYGESLKKCLGKEVARVESFHGFTYGCDFDQAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPBGKVIKANKKAPENFVGDAHYHGWTHASSLSRSGESIFSSLAGNALLPREGAGL 240  
DB 181 ELVGPBGKVIKANKKAPENFVGDAHYHGWTHASSLSRSGESIFSSLAGNALLPREGAGL 240  
QY 241 QMTSKYSGSGMGLMDGYSGVSHSADLVPELMARFGAKQERLNEIGDVRARIYRSHLNCV 300  
DB 241 QMTSKYSGSGMGLMDGYSGVSHSADLVPELMARFGAKQERLNEIGDVRARIYRSHLNCV 300  
QY 301 FPNNSMLTCSGVEFKYKWNPIDANTTEWYTAIVEKMPEDLKRRLADSVQRTGPGAFWES 360  
DB 301 FPNNSMLTCSGVEFKYKWNPIDANTTEWYTAIVEKMPEDLKRRLADSVQRTGPGAFWES 360  
QY 361 DDNDMMETASQNGKKYQSDLSNLGFGEDYGDAYPGVGSALGETSYRGFYRAY 420  
DB 361 DDNDMMETASQNGKKYQSDLSNLGFGEDYGDAYPGVGSALGETSYRGFYRAY 420  
QY 421 QAHVSSNMMAEFHASTWHTELTKTTDR 449  
DB 421 QAHVSSNMMAEFHASTWHTELTKTTDR 449

RESULT 4  
AAB12583  
ID AAB12583 standard; Protein: 449 AA.  
XX  
AC AAB12583;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352W protein sequence SEQ ID NO:58.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.

```
XX 26-OCT-1998; 98US-0105575.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX
XX Parales R, Gibson D, Resnick S, Lee K;
XX WPI; 2000-452174/39.
XX N-PSDB; AAA65376.
XX
XX Novel naphthalene dioxygenase mutant having a specific amino acid
XX substitution for preparing chiral diols for use in the polymer, resin,
XX pharmaceutical or rubber industry and for carrying out bioremediation
XX
XX
XX Example 7; Page 142-144; 151pp; English.
XX
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO
XX related complex (I) comprising several polypeptides which contain an
XX alpha subunit that contains substituted amino acids at specific
XX positions. The polypeptides and host cells are useful for preparing
XX (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
XX dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2,2-
XX dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
XX involves contacting them with naphthalene, biphenyl, phenanthrene,
XX indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
XX the host cells are also useful for preparing 1,2-dihydroxy-1,2-
XX dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
XX involves contacting them with phenanthrene. The polypeptides and the
XX host cells are also used in bioremediation in which they oxidize an
XX aromatic compound such as indene, 1,2-dihydronaphthalene,
XX benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,
XX acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
XX benzophenone, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a
XX corresponding dihydro dihydroxy compound. The polypeptides and the host
XX cells are useful for preparing chiral diols for use in the polymer,
XX resin, pharmaceutical or rubber industry. The present sequence represents
XX a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
XX exemplification of the present invention.
XX
XX Sequence 449 AA:
XX
XX Query Match 99.78; Score 2402; DB 21; Length 449;
XX Best Local Similarity 99.88; Pred. No. 8.2e-224;
XX Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MNVNNKILVSESLGSLKHLHGDELFQHEKLTFFARNWMLFTLTDLSLPARGDVTAKMG 60
DB 1 MNVNNKILVSESLGSLKHLHGDELFQHEKLTFFARNWMLFTLTDLSLPARGDVTAKMG 60
QY 61 IDEVIVSRONGSIRAFNLVNCRRHKGTLVSEAGNAGFCVSYGNGFGSGNGELQSVPE 120
DB 61 IDEVIVSRONGSIRAFNLVNCRRHKGTLVSEAGNAGFCVSYGNGFGSGNGELQSVPE 120
QY 121 KDLVGSLLNKKCLGLKEVARVESFHGTYGCFDQEARPLMDYLCDAAWYLEPMFKHSGGL 180
DB 121 KDLVGSLLNKKCLGLKEVARVESFHGTYGCFDQEARPLMDYLCDAAWYLEPMFKHSGGL 180
QY 122 KDLVGSLLNKKCLGLKEVARVESFHGTYGCFDQEARPLMDYLCDAAWYLEPMFKHSGGL 180
DB 122 KDLVGSLLNKKCLGLKEVARVESFHGTYGCFDQEARPLMDYLCDAAWYLEPMFKHSGGL 180
QY 181 ELVGPGRKVIYIKANWKPAPNFYGDAAHYVGMTAHSLSRSGSTSSLAGNALRPEAGL 240
DB 181 ELVGPGRKVIYIKANWKPAPNFYGDAAHYVGMTAHSLSRSGSTSSLAGNALRPEAGL 240
QY 241 QMTSKYSGMGVLMDSYGSVHSDILVPELMAFGAKOERLKEIGDVRARIYRSHLNTCV 300
DB 241 QMTSKYSGMGVLMDSYGSVHSDILVPELMAFGAKOERLKEIGDVRARIYRSHLNTCV 300
QY 301 FPNNSMLTSGGVFVKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTGPAFWES 360
DB 301 FPNNSMLTSGGVFVKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTGPAFWES 360
QY 361 DQNDNMTASONGKRYOSRSDLSNLGFGEDVYGDAAVYGVGKSAIGETSYRGFYRAY 420
DB 361 DQNDNMTASONGKRYOSRSDLSNLGFGEDVYGDAAVYGVGKSAIGETSYRGFYRAY 420
```

```
QY 421 QAHVSSNMAEFEHASSWTWTELTCTDR 449.
DB 421 QAHVSSNMAEFEHASSWTWTELTCTDR 449
RESULT 5
AAB12565
ID AAB12565 standard; Protein; 449 AA.
XX
XX AAB12565;
XX
XX 09-NOV-2000 (first entry)
XX
XX Naphthalene dioxygenase mutant F352V protein sequence SEQ ID NO:2.
XX
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
XX inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
XX polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX
XX OS Pseudomonas sp.
XX Synthetic.
XX
XX W0200037480-A1.
XX
XX 29-JUN-2000.
XX
XX 26-OCT-1999; 99WO-US25079.
XX
XX 26-OCT-1998; 98US-0105575.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX
XX Parales R, Gibson D, Resnick S, Lee K;
XX WPI; 2000-452174/39.
XX N-PSDB; AAA65376.
XX
XX Novel naphthalene dioxygenase mutant having a specific amino acid
XX substitution for preparing chiral diols for use in the polymer, resin,
XX pharmaceutical or rubber industry and for carrying out bioremediation
XX
XX
XX Claim 9; Page 58-60; 151pp; English.
XX
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO
XX related complex (I) comprising several polypeptides which contain an
XX alpha subunit that contains substituted amino acids at specific
XX positions. The polypeptides and host cells are useful for preparing
XX (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
XX dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2,2-
XX dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
XX involves contacting them with naphthalene, biphenyl, phenanthrene,
XX indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
XX the host cells are also useful for preparing 1,2-dihydroxy-1,2-
XX dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
XX involves contacting them with phenanthrene. The polypeptides and the
XX host cells are also used in bioremediation in which they oxidize an
XX aromatic compound such as indene, 1,2-dihydronaphthalene,
XX benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,
XX acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
XX benzophenone, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a
XX corresponding dihydro dihydroxy compound. The polypeptides and the host
XX cells are useful for preparing chiral diols for use in the polymer,
XX resin, pharmaceutical or rubber industry. The present sequence represents
XX a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
XX exemplification of the present invention.
XX
XX Sequence 449 AA:
XX
XX Query Match 99.68; Score 2401; DB 21; Length 449;
XX Best Local Similarity 99.88; Pred. No. 1e-223;
XX Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MNYNNKILVSESGLSQKHILHGDEELFOHELKTIFFARNMLFLTHDSLIPAGDYVYAKMG 60
D 1 MNYNNKILVSESGLSQKHILHGDEELFOHELKTIFFARNMLFLTHDSLIPAGDYVYAKMG 60
QY 61 IDEVIVSRQNDGSIIRAFLANVCRHRGKTLVSEAGNAKGFCVSYHGMGFGSNGELQSVPE 120
D 61 IDEVIVSRQNDGSIIRAFLANVCRHRGKTLVSEAGNAKGFCVSYHGMGFGSNGELQSVPE 120
QY 121 KDLYGESLNNKCKGLKEVAVESFHGFIYGCFOEAPPLMDYIGDAWYIEPFFKHSGL 180
D 121 KDLYGESLNNKCKGLKEVAVESFHGFIYGCFOEAPPLMDYIGDAWYIEPFFKHSGL 180
QY 181 ELVGPPEKVVIRKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPPREGAGL 240
D 181 ELVGPPEKVVIRKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPPREGAGL 240
QY 241 QMTSKYSGMGVLMDSYGVHSADLVEPMLAFGAKOERLNKEIGVRRARIYRSHLNCY 300
D 241 QMTSKYSGMGVLMDSYGVHSADLVEPMLAFGAKOERLNKEIGVRRARIYRSHLNCY 300
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVYAIYEKDMPEDLKRRLADSVORTGPGPWES 360
D 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVYAIYEKDMPEDLKRRLADSVORTGPGPWES 360
QY 361 DDNDNMETASONGKRYQSRSDLSNLGFGEDYVGDVYGVVGSALGETSYRGFYRAY 420
D 361 DDNDNMETASONGKRYQSRSDLSNLGFGEDYVGDVYGVVGSALGETSYRGFYRAY 420
QY 421 QAHVSSNNAAEFHASTWHTLTCTDR 449
D 421 QAHVSSNNAAEFHASTWHTLTCTDR 449

RESULT 6
AAB12566
ID AAB12566 standard; Protein: 449 AA.
XX
AC AAB12566;
XX
DT 09-NOV-2000 (first entry)
DE NDO related complex alpha subunit protein sequence SEQ ID NO:14.
XX
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX
OS Pseudomonas sp.
XX
SY Synthetic.
XX
PN MO200037480-A1.
XX
PD 29-JUN-2000.
XX
PF 26-OCT-1999; 99WO-US25079.
XX
PR 26-OCT-1998; 98US-0105575.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PI Parales R, Gibson D, Resnick S, Lee K;
XX
DR WPI: 2000-452174/39.
XX
DR N-PSDB; AAA65340.
XX
PT Novel naphthalene dioxygenase mutant having a specific amino acid
substitution for preparing chiral diols for use in the polymer, resin,
pharmaceutical or rubber industry and for carrying out bioremediation
PT
XX
PS Claim 13; Page 99-100; 151pp; English.
XX
```

```
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO
CC related complex (1) comprising several polypeptides which contain an
CC alpha subunit that contains substituted amino acids at specific
CC positions. The polypeptides and host cells are useful for preparing
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
CC involves contacting them with naphthalene, biphenyl, phenanthrene,
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
CC involves contacting them with phenanthrene. The polypeptides and the
CC host cells are also used in bioremediation in which they oxidise an
CC aromatic compound such as indene, 1,2-dihydronaphthalene,
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzoc(1,4)dioxan,
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
CC dibenzochlorophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a
CC corresponding dihydro dihydro compound. The polypeptides and the host
CC cells are useful for preparing chiral diols for use in the polymer,
CC resin, pharmaceutical or rubber industry. The present sequence represents
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
CC exemplification of the present invention.
XX
SQ Sequence 449 AA:
XX
Query Match 99.6%; Score 2401; DB 21; Length 449;
Best Local Similarity 99.8%; Pred. No. 1e-223; 1; Indels 0; Gaps 0;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MNYNNKILVSESGLSQKHILHGDEELFOHELKTIFFARNMLFLTHDSLIPAGDYVYAKMG 60
D 1 MNYNNKILVSESGLSQKHILHGDEELFOHELKTIFFARNMLFLTHDSLIPAGDYVYAKMG 60
QY 61 IDEVIVSRQNDGSIIRAFLANVCRHRGKTLVSEAGNAKGFCVSYHGMGFGSNGELQSVPE 120
D 61 IDEVIVSRQNDGSIIRAFLANVCRHRGKTLVSEAGNAKGFCVSYHGMGFGSNGELQSVPE 120
QY 121 KDLYGESLNNKCKGLKEVAVESFHGFIYGCFOEAPPLMDYIGDAWYIEPFFKHSGL 180
D 121 KDLYGESLNNKCKGLKEVAVESFHGFIYGCFOEAPPLMDYIGDAWYIEPFFKHSGL 180
QY 181 ELVGPPEKVVIRKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPPREGAGL 240
D 181 ELVGPPEKVVIRKANMKAPAEFNGDAVHVGWTHASSLRSGESIFSSLAGNAALPPREGAGL 240
QY 241 QMTSKYSGMGVLMDSYGVHSADLVEPMLAFGAKOERLNKEIGVRRARIYRSHLNCY 300
D 241 QMTSKYSGMGVLMDSYGVHSADLVEPMLAFGAKOERLNKEIGVRRARIYRSHLNCY 300
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVYAIYEKDMPEDLKRRLADSVORTGPGPWES 360
D 301 FPNNSMLTCSGVFKVWNPIDANTTEWTVYAIYEKDMPEDLKRRLADSVORTGPGPWES 360
QY 361 DDNDNMETASONGKRYQSRSDLSNLGFGEDYVGDVYGVVGSALGETSYRGFYRAY 420
D 361 DDNDNMETASONGKRYQSRSDLSNLGFGEDYVGDVYGVVGSALGETSYRGFYRAY 420
QY 421 QAHVSSNNAAEFHASTWHTLTCTDR 449
D 421 QAHVSSNNAAEFHASTWHTLTCTDR 449

RESULT 7
AAB12567
ID AAB12567 standard; Protein: 449 AA.
XX
AC AAB12567;
XX
DT 09-NOV-2000 (first entry)
DE NDO related complex alpha subunit protein sequence SEQ ID NO:15.
XX
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
XX
```



CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC naphthalene dioxygenase (NDO) from the *Pseudomonas* sp. strain NCIB  
 CC 9816-4, which is used in the exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.6%; Score 2401; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 1e-223;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNTNNKILVSESGISQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYVAKMG 60  
 DB 1 MNTNNKILVSESGISQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYVAKMG 60  
 QY 61 IDEVIYSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNKGFCVSHGWFSGNGLQSVPE 120  
 DB 61 IDEVIYSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNKGFCVSHGWFSGNGLQSVPE 120  
 QY 121 KDLYGESLNKKCLGKLEVARVESFHGFIYGCFOQEARPLMDYLGDAAWYLEPMFKHSGCL 180  
 DB 121 KDLYGESLNKKCLGKLEVARVESFHGFIYGCFOQEARPLMDYLGDAAWYLEPMFKHSGCL 180  
 QY 181 ELVGPPEKVVIRKANKKAPAEFNVGDYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240  
 DB 181 ELVGPPEKVVIRKANKKAPAEFNVGDYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240  
 QY 241 QMTSKYSGGCVLMDYSGVHSADLPELMAFGAEOERLNKEIGVRRARIYSHLNCY 300  
 DB 241 QMTSKYSGGCVLMDYSGVHSADLPELMAFGAEOERLNKEIGVRRARIYSHLNCY 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKDMPEDLKRLADSVORTGPGAFWES 360  
 DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKDMPEDLKRLADSVORTGPGAFWES 360  
 QY 361 DDNDNMETASQNKKTQSRSDLSNLGFGEDYGGDAVPGVYKSAIGETSRGYRAY 420  
 DB 361 DDNDNMETASQNKKTQSRSDLSNLGFGEDYGGDAVPGVYKSAIGETSRGYRAY 420  
 QY 421 QAHVSSNNAEFPHASSTWHTELTKTTDR 449  
 DB 421 QAHVSSNNAEFPHASSTWHTELTKTTDR 449

RESULT 9  
 AAB12581  
 ID AAB12581 standard; Protein; 449 AA.  
 XX  
 AC AAB12581;  
 XX  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE Naphthalene dioxygenase mutant F352L protein sequence SEQ ID NO:35.  
 XX  
 KW *Pseudomonas* sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 XX  
 OS *Pseudomonas* sp.  
 OS Synthetic.  
 OS  
 PN WO200037480-A1.  
 PD 29-JUN-2000.  
 PF 26-OCT-1999; 99MO-US25079.  
 PR 26-OCT-1998; 98US-0105575.  
 PA (IOWA ) UNIV IOWA RES FOUND.  
 PI Parales R, Gibson D, Resnick S, Lee K;  
 XX

DR WPI: 2000-452174/39.  
 DR N-PSDB: AAA53355.

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT  
 PS Claim 9; Page 131-133; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (1) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidize an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzophenone, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.6%; Score 2400; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-223;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNTNNKILVSESGISQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYVAKMG 60  
 DB 1 MNTNNKILVSESGISQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYVAKMG 60  
 QY 61 IDEVIYSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNKGFCVSHGWFSGNGLQSVPE 120  
 DB 61 IDEVIYSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNKGFCVSHGWFSGNGLQSVPE 120  
 QY 121 KDLYGESLNKKCLGKLEVARVESFHGFIYGCFOQEARPLMDYLGDAAWYLEPMFKHSGCL 180  
 DB 121 KDLYGESLNKKCLGKLEVARVESFHGFIYGCFOQEARPLMDYLGDAAWYLEPMFKHSGCL 180  
 QY 181 ELVGPPEKVVIRKANKKAPAEFNVGDYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240  
 DB 181 ELVGPPEKVVIRKANKKAPAEFNVGDYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240  
 QY 241 QMTSKYSGGCVLMDYSGVHSADLPELMAFGAEOERLNKEIGVRRARIYSHLNCY 300  
 DB 241 QMTSKYSGGCVLMDYSGVHSADLPELMAFGAEOERLNKEIGVRRARIYSHLNCY 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKDMPEDLKRLADSVORTGPGAFWES 360  
 DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKDMPEDLKRLADSVORTGPGAFWES 360  
 QY 361 DDNDNMETASQNKKTQSRSDLSNLGFGEDYGGDAVPGVYKSAIGETSRGYRAY 420  
 DB 361 DDNDNMETASQNKKTQSRSDLSNLGFGEDYGGDAVPGVYKSAIGETSRGYRAY 420  
 QY 421 QAHVSSNNAEFPHASSTWHTELTKTTDR 449  
 DB 421 QAHVSSNNAEFPHASSTWHTELTKTTDR 449

RESULT 10

XX	AAAB12582
ID	AAAB12582 standard; Protein; 449 AA.
AC	AAAB12582;
XX	
XX	AAAB12582;
DT	09-NOV-2000 (first entry)
DE	Naphthalene dioxygenase mutant F352I protein sequence SEQ ID NO:36.
XX	
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
RW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KV	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
OS	Pseudomonas sp.
OS	Synthetic.
PN	MO200037480-AI.
PD	29-JUN-2000.
XX	
PF	26-OCT-1999; 99WO-US25079.
XX	
PR	26-OCT-1998; 98US-0105575.
XX	
PA	(IOWA ) UNIV IOWA RES FOUND.
PI	Parales R, Gibson D, Resnick S, Lee K;
XX	
DR	WPI: 2000-452174/39.
DR	N-PSDB: AAA65356.
PT	Novel naphthalene dioxygenase mutant having a specific amino acid
PT	substitution for preparing chiral diols for use in the polymer, resin,
PT	pharmaceutical or rubber industry and for carrying out bioremediation
PS	-
XX	
PS	Claim 9; Page 133-134; 151pp; English.
XX	
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO
CC	related complex (I) comprising several polypeptides which contain an
CC	alpha subunit that contains substituted amino acids at specific
CC	positions. The polypeptides and host cells are useful for preparing
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-
CC	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-
CC	dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
CC	involves contacting them with phenanthrene. The polypeptides and the
CC	host cells are also used in bioremediation in which they oxidise an
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,
CC	benzocyclohept-1-ene, anthracene, phenanthrene, dibenzofuran, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
CC	dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a
CC	corresponding dlydro dihydro compound. The polypeptides and the host
CC	cells are useful for preparing chiral diols for use in the polymer,
CC	resin, pharmaceutical or rubber industry. The present sequence represents
CC	a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
CC	exemplification of the present invention.
XX	
XX	Sequence 449 AA:
XX	5Q
Query Match	99.6%; Score 2400; DB 21; Length 449;
Best Local Similarity	99.8%; Pred. No. 1.3e+223;
Matches 448; Conservative 0;	Mismatches 1; Indels 0; Gaps 0;
OY	1 MNYNNKILVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPDDYATKMG 60
Db	1 MNYNKKLIVSESGLSQKHILHGDDELFOHELKTIFARNMLFLTHDSLIPAPDDYATKMG 60
OY	61 IDEVTVSRQNDGSIARFLAVCRHRGKITLVSVAGNAKGVCSTHGCGFSGNELGVPE 120

Db	61	IDEVTVSRQNDGSTRAPFLNCRHNGKTLVSYEAGNAKGFVCSITHGMGFSGNCELOSYPE	120
Qy	121	KDLVEESLNKKCKLGEKVARVESEHGEIYGCFDGEAPPLMDYLGDAAWYLEPMFKHSGSL	180
Db	121	KDLVESLNKKCKLGEKVARVESPHGFYGCDFGEAPPLMDYLGDAAWYLEPMFKHSGSL	180
Qy	181	ELVGPFGVYIKAMKKAPENFVDAYHVGWTHASSLRGSESTFFSSLAGNAALPEBGAGL	240
Db	181	ELVGPFGVYIKAMKKAPENFVDADYHVGWTHASSLRGSESTFFSSLAGNAALPEBGAGL	240
Qy	241	QMTSKYSGMGVLDGSGVSHADLYPELIAFGAKOERLNKTIIGVRRARITRSHLNCIV	300
Db	241	QMTSKYSGMGVLDGSGVSHADLYPELIAFGAKOERLNKTIIGVRRARITRSHLNCIV	300
Qy	301	FPNNMMLTCSGVFVFWMPIDANTTEVWTVYAIVEKMDPEDKRLRLADSVORTGGPAGFWS	360
Db	301	FPNNMMLTCSGVFVFWMPIDANTTEVWTVYAIVEKMDPEDKRLRLADSVORTGGPAGFWS	360
Qy	361	DDNDNMEFASONGKRRKYOSRSDLSLNGFGEEDVYGDAVYPGVVGSKAIGETSYRGCFYRAY	420
Db	361	DDNDNMEFASONGKRRKYOSRSDLSLNGFGEEDVYGDAVYPGVVGSKAIGETSYRGCFYRAY	420
Qy	421	QAHVSSNMWAEFEHASSHTWHTLTKTTDR	449
Db	421	QAHVSSNMWAEFEHASSHTWHTLTKTTDR	449
RESULT 11			
AAB12584			
ID	AAB12584	standard; Protein; 448 AA.	
AC	AAB12584;		
XX			
DT	09-NOV-2000	(first entry)	
XX			
DE	Naphthalene dioxygenase mutant F352Y protein sequence SEQ ID NO:59.		
XX			
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;		
KM	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
KM	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
XX			
OS	Pseudomonas sp.		
OS	Synthetic.		
PN	WO200037480-A1.		
PD	29-JUN-2000.		
XX			
PF	26-OCT-1999; 99WO-US25079.		
XX			
PR	26-OCT-1998; 98US-0105575.		
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
PI	Parales R, Gibson D, Resnick S, Lee K;		
DR	WP1: 2000-452174/39.		
DR	N-PSDB; AAA65377.		
XX			
PT	Novel naphthalene dioxygenase mutant having a specific amino acid		
PT	substitution for preparing chiral diols for use in the polymer, resin,		
PT	pharmaceutical or rubber industry and for carrying out bioremediation		
XX			
PS	Example 7; Page 144-145; 151pp; English.		
XX			
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO		
CC	related complex (I) comprising several polypeptides which contain an		
CC	alpha subunit that contains substituted amino acids at specific		
CC	positions. The polypeptides and host cells are useful for preparing		
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-		
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-		
CC	dihydroxindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which		



CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 448 AA:

Query Match 99.4%; Score 2396; DB 21; Length 448;  
 Best Local Similarity 99.8%; Pred. No. 3,1e-223;  
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVNNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFTLHSLIPAPGDYVAKKG 60  
 DB 1 MNVNNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFTLHSLIPAPGDYVAKKG 60  
 QY 61 IDEVYSRONDSIRAFNLVNCRRHRTILVSVAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 DB 61 IDEVYSRONDSIRAFNLVNCRRHRTILVSVAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 QY 121 KDLVGSLENNKCLGLKEVARVESFHGFIYGCDOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 DB 121 KDLVGSLENNKCLGLKEVARVESFHGFIYGCDOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPGRKVVITANKKAPENFVGDAVHYGWTTHASSLRGSESTFSLAANAALPPGAGL 240  
 DB 181 ELVGPGRKVVITANKKAPENFVGDAVHYGWTTHASSLRGSESTFSLAANAALPPGAGL 240  
 QY 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKOEKRLNEIGVRRARIYRSHLNTCV 300  
 DB 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKOEKRLNEIGVRRARIYRSHLNTCV 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDMPEDLKRLADSVOGTGPGAGFWS 360  
 DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDMPEDLKRLADSVOGTGPGAGFWS 360  
 QY 361 DDDNNMETASQNGKKYQSDLSNLGFGEDVYGDVAPGVYKSAIGETSRGYRAY 420  
 DB 361 DDDNNMETASQNGKKYQSDLSNLGFGEDVYGDVAPGVYKSAIGETSRGYRAY 420  
 QY 421 QAHVSSSNMAEEFHASSTWHTELTKTTD 448  
 DB 421 QAHVSSSNMAEEFHASSTWHTELTKTTD 448

RESULT 12  
 AAB12568  
 ID AAB12568 standard; Protein: 449 AA.  
 XX  
 AC AAB12568;  
 XX  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE NDO related complex alpha subunit protein sequence SEQ ID NO:16.  
 XX  
 KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 XX  
 OS Pseudomonas sp.  
 XX Synthetic.  
 OS  
 PN WO200037480-A1.

XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 PF 26-OCT-1998; 98US-0105575.  
 PR (IOWA) UNIV IOWA RES FOUND.  
 PA  
 PI Parles R, Gibson D, Resnick S, Lee K;  
 DR WPI: 2000-452174/39.  
 DR N-PSDB: AAA65342.  
 PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 .

Claim 13; Page 102-103; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (1) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.4%; Score 2396; DB 21; Length 449;  
 Best Local Similarity 99.6%; Pred. No. 3,1e-223;  
 Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNVNNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFTLHSLIPAPGDYVAKKG 60  
 DB 1 MNVNNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFTLHSLIPAPGDYVAKKG 60  
 QY 61 IDEVYSRONDSIRAFNLVNCRRHRTILVSVAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 DB 61 IDEVYSRONDSIRAFNLVNCRRHRTILVSVAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 QY 121 KDLVGSLENNKCLGLKEVARVESFHGFIYGCDOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 DB 121 KDLVGSLENNKCLGLKEVARVESFHGFIYGCDOEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPGRKVVITANKKAPENFVGDAVHYGWTTHASSLRGSESTFSLAANAALPPGAGL 240  
 DB 181 ELVGPGRKVVITANKKAPENFVGDAVHYGWTTHASSLRGSESTFSLAANAALPPGAGL 240  
 QY 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKOEKRLNEIGVRRARIYRSHLNTCV 300  
 DB 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKOEKRLNEIGVRRARIYRSHLNTCV 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDMPEDLKRLADSVOGTGPGAGFWS 360  
 DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDMPEDLKRLADSVOGTGPGAGFWS 360

QY 361 DDNDNMTASONGKRRYOSRSDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 DDNDNMTASONGKRRYOSRSDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFHASSSTWHTELTKTTDR 449  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 QAHVSSSNMAEFHASSSTWHTELTKTTDR 449

## RESULT 13

AAB12569 standard; Protein: 449 AA.

AC AAB12569;  
 XX  
 DT 09-NOV-2000 (first entry)

DE NDO related complex alpha subunit protein sequence SEQ ID NO:17.

KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.

XX Pseudomonas sp.  
 OS Synthetic.

PN WO200037480-A1.

PD 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB: AAA65343.

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation

PS Claim 13; Page 104-105; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (i) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA;  
 SQ

Query Match 99.2%; Score 2391; DB 21; Length 449;  
 Best Local Similarity 99.3%; Pred. No. 9.6e-223;  
 Matches 446; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNYNKILVSESGLSQKHLIHGDEELFOHELKTIIFARNWMLFTHDLSIPAPGYVAKMG 60  
 Db 1 MNYNKILVSESGLSQKHLIHGDEELFOHELKTIIFARNWMLFTHDLSIPAPGYVAKMG 60  
 QY 61 IDEVIVSRQNDGSTRAPLNVCRHRGKTLIVSEAGNAKGFVCSYHGKFGSNGELQSVPE 120  
 Db 61 IDEVIVSRQNDGSTRAPLNVCRHRGKTLIVSEAGNAKGFVCSYHGKFGSNGELQSVPE 120  
 QY 121 KDLYGESLNKKCLGKEVARVESFHGFIYCGFDEAPPLMDYLGDADAWYLEPMFKHSGGL 180  
 Db 121 KDLYGESLNKKCLGKEVARVESFHGFIYCGFDEAPPLMDYLGDADAWYLEPMFKHSGGL 180  
 QY 181 ELVGPFGKRVVIRKAWKKAFAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPGAGL 240  
 Db 181 ELVGPFGKRVVIRKAWKKAFAENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPGAGL 240  
 QY 241 QMTSKYSGMGVLDGTSVGSADLVEPLMAFGAKOEKRLKEIGVRAITYSHLNCTY 300  
 Db 241 QMTSKYSGMGVLDGTSVGSADLVEPLMAFGAKOEKRLKEIGVRAITYSHLNCTY 300  
 QY 301 FPNNSMILTCGCVFRVWNPIDANTTEVMTVAIVEKMPEDLKRRLADSVORTGGPAGFWS 360  
 Db 301 FPNNSMILTCGCVFRVWNPIDANTTEVMTVAIVEKMPEDLKRRLADSVORTGGPAGFWS 360  
 QY 361 DDNDNMTASONGKRRYOSRSDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 Db 361 DDNDNMTASONGKRRYOSRSDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFHASSSTWHTELTKTTDR 449  
 Db 421 QAHVSSSNMAEFHASSSTWHTELTKTTDR 449

## RESULT 14

AAB12570 standard; Protein: 449 AA.

AC AAB12570;

DT 09-NOV-2000 (first entry)

DE NDO related complex alpha subunit protein sequence SEQ ID NO:18.

KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KW polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS Pseudomonas sp.  
 OS Synthetic.

PN WO200037480-A1.

PD 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB: AAA65344.

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation

PT -  
XX  
PS  
XX  
XX  
Claim 13; Page 105-107; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (1) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 449 AA:  
Query Match 97.2%; Score 2343; DB 21; Length 449;  
Best Local Similarity 96.7%; Pred. No. 4,3e-218;  
Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MMYNNKIIVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLRAPGDYVYAKMG 60  
DB 1 MMYNNKIIVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLRAPGDYVYAKMG 60  
QY 61 IDEVIVSRONDGSIRAFNLVNCRRHGRKTLVSEVAGNAKGFCVSYHGNGFSGNGLQSVPE 120  
DB 61 IDEVIVSRONDGSIRAFNLVNCRRHGRKTLVSEVAGNAKGFCVSYHGNGFSGNGLQSVPE 120  
QY 121 KDIYGSLSNKKCLGKLEVARVESFHGTTCGPOEAPRLMDYGDAAWYLEPFRKISGGL 180  
DB 121 KDIYGSLSNKKCLGKLEVARVESFHGTTCGPOEAPRLMDYGDAAWYLEPFRKISGGL 180  
QY 121 KELYGSLSNKKCLGKLEVARVESFHGTTCGPOEAPRLMDYGDAAWYLEPFRKISGGL 180  
DB 121 KELYGSLSNKKCLGKLEVARVESFHGTTCGPOEAPRLMDYGDAAWYLEPFRKISGGL 180  
QY 181 ELVGPGRKVIYKANKAPENFVGDAVHYGWTASSLRSGESFSSLAGMAALPPRGAGL 240  
DB 181 ELVGPGRKVIYKANKAPENFVGDAVHYGWTASSLRSGESFSSLAGMAALPPRGAGL 240  
QY 241 QMTSKYSGMGLMDGYSGVHSADLVPBELMAFGAKOERLNKEIGDVRARIYSHLNCIV 300  
DB 241 QMTSKYSGMGLMDGYSGVHSADLVPBELMAFGAKOERLNKEIGDVRARIYSHLNCIV 300  
QY 301 FPNNSMLTSGVFKVWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTGGPAGFWES 360  
DB 301 FPNNSMLTSGVFKVWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTGGPAGFWES 360  
QY 361 DDNDNNEFASONGKYYOSRSDLSNLGFGEDYVSDAVYPGVYKSAIETSRGGRYAY 420  
DB 361 DDNDNNEFASONGKYYOSRSDLSNLGFGEDYVSDAVYPGVYKSAIETSRGGRYAY 420  
QY 421 QAHVSSSNAEPEHASTWHTLTKTTDR 449  
DB 421 QAHVSSSNAEPEHASTWHTLTKTTDR 449  
RESULT 15  
AAB12571  
ID AAB12571 standard; Protein: 449 AA.  
XX  
AC AAB12571;  
XX  
DT 09-NOV-2000 (first entry)

XX  
DE NDO related complex alpha subunit protein sequence SEQ ID NO:19.  
XX  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
OS Pseudomonas sp.  
OS Synthetic.  
FH Key location/Qualifiers  
FT Misc-difference 35  
FT  
PN MO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
XX 26-OCT-1999; 99MO-US25079.  
PF  
XX 26-OCT-1998; 98US-0105575.  
PR  
XX (IOWA ) UNIV IOWA RES FOUNO.  
PA  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
DR N-PSDB; AAB65345.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
PT  
PS  
XX  
XX  
Claim 13; Page 107-109; 151pp; English.  
CC  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (1) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 449 AA:  
Query Match 96.1%; Score 2315; DB 21; Length 449;  
Best Local Similarity 95.5%; Pred. No. 2,2e-215;  
Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;  
QY 1 MMYNNKIIVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLRAPGDYVYAKMG 60  
DB 1 MMYNNKIIVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLRAPGDYVYAKMG 60  
QY 61 IDEVIVSRONDGSIRAFNLVNCRRHGRKTLVSEVAGNAKGFCVSYHGNGFSGNGLQSVPE 120  
DB 61 IDEVIVSRONDGSIRAFNLVNCRRHGRKTLVSEVAGNAKGFCVSYHGNGFSGNGLQSVPE 120

Qy	121	KDIYGESENKCKOLGKEVARARESHGFIYGCFODEAPLMDYICGDAMWLEPMFKHSGGL	180
Db	121	KELGESENKCKOLGKEVARARESHGFIYGCFODEAPSLMDYICGDAMWLEPIFKHSGGL	180
Qy	181	ELVGPBGKVVYIKANKKAPAEWVGDAYVGTTHASSLSRSGESTFSSLAGNAALPPBGAGL	240
Db	181	ELVGPBGKVVYIKANKKAPAEWVGDAYVGTTHASSLRTGSESTFSSLAGNAVLPPBGAGL	240
Qy	241	QMTSKYSGSGMGVLMDGYSGVSHADLVPELMAFGAKOEBRLKEIGDVRARITYRSHLNCY	300
Db	241	QMTSKYSGSGMGVLMDGYSGVSHADLVPELMAFGAKOEBRLKEIGDVPARITYRSHLNCY	300
Qy	301	FPNNSMLTCSGVFYKVMNDIDANTTEWMTYAIVEKDMPEDLKRLADSVOQRTGPPAGFWS	360
Db	301	FPNNSVLTCSGVFYKVMNIDANTTEWMTYAIVEKDMPEDLKRLADAVQRTGPPAGFWS	360
Qy	361	DDNDMELTASQNGKKYQSRDSDLSNLGFGGDVYGDVAYPGVYKSAIGETSTRGFRAY	420
Db	361	DDNDMELTASQNGKKYQSRDSDLSNLGFGKDVYGDVAYPGVYKSAIGETSTRGFRAY	420
Qy	421	QAHVSSNMWAEFEHASTFWHLETKTTDR	449
Db	421	QAHVSSNMWAEEDASTFWHLETKTTDR	449

```
Search completed: January 27, 2003, 08:58:47
Job time : 31.1667 secs
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 09:02:32 : Search time 7 Seconds  
(without alignments)  
1294.310 Million cell updates/sec

Title: US-09-843-250-33

Perfect score: 2408

Sequence: 1 MNYNKILVSESGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC01\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PC01\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	735	30.5	453	10 US-09-815-242-10253	Sequence 10253, A
2	375.5	15.6	385	10 US-09-815-242-11692	Sequence 11692, A
3	375	15.6	490	9 US-09-738-626-6140	Sequence 6140, Ap
4	336.5	14.0	424	10 US-08-815-242-5097	Sequence 5097, Ap
5	185	7.7	35	10 US-09-776-490-19	Sequence 19, Appl
6	185	7.7	35	10 US-09-776-491-19	Sequence 19, Appl
7	174	7.2	35	10 US-09-776-490-20	Sequence 20, Appl
8	174	7.2	35	10 US-09-776-491-20	Sequence 20, Appl
9	168	7.0	35	10 US-09-776-490-21	Sequence 21, Appl
10	168	7.0	35	10 US-09-776-491-21	Sequence 21, Appl
11	118.5	4.9	354	1 US-08-976-063C-4	Sequence 4, Appl1
12	110	4.6	35	10 US-09-776-490-12	Sequence 12, Appl
13	110	4.6	35	10 US-09-776-491-12	Sequence 12, Appl
14	109	4.5	35	10 US-09-776-490-14	Sequence 14, Appl
15	109	4.5	35	10 US-09-776-491-14	Sequence 14, Appl
16	108	4.5	35	10 US-09-776-490-13	Sequence 13, Appl
17	108	4.5	35	10 US-09-776-490-15	Sequence 15, Appl
18	108	4.5	35	10 US-09-776-491-13	Sequence 13, Appl
19	108	4.5	35	10 US-09-776-491-15	Sequence 15, Appl

20	107.5	4.5	548	9 US-10-047-542-78	Sequence 78, Appl
21	106	4.4	35	10 US-09-776-490-9	Sequence 9, Appl1
22	106	4.4	35	10 US-09-776-491-9	Sequence 9, Appl1
23	105.5	4.4	951	9 US-09-924-097-15	Sequence 15, Appl
24	104	4.3	35	10 US-09-776-490-18	Sequence 18, Appl
25	104	4.3	35	10 US-09-776-491-18	Sequence 18, Appl
26	103	4.3	35	10 US-09-776-490-11	Sequence 11, Appl
27	103	4.3	35	10 US-09-776-491-11	Sequence 11, Appl
28	102	4.2	35	10 US-09-776-490-8	Sequence 8, Appl1
29	102	4.2	35	10 US-09-776-490-10	Sequence 10, Appl
30	102	4.2	35	10 US-09-776-491-8	Sequence 8, Appl1
31	102	4.2	35	10 US-09-776-491-10	Sequence 10, Appl
32	101.5	4.2	376	9 US-09-738-626-6115	Sequence 6115, Ap
33	101	4.2	17	10 US-09-776-490-45	Sequence 45, Appl
34	101	4.2	17	10 US-09-776-490-45	Sequence 45, Appl
35	101	4.2	17	10 US-09-776-491-44	Sequence 44, Appl
36	101	4.2	17	10 US-09-776-491-45	Sequence 45, Appl
37	100	4.2	35	10 US-09-776-490-17	Sequence 17, Appl
38	100	4.2	35	10 US-09-776-491-17	Sequence 17, Appl
39	97	4.0	3069	9 US-09-712-363-246	Sequence 246, App
40	95	3.9	17	10 US-09-776-490-46	Sequence 46, Appl
41	95	3.9	17	10 US-09-776-491-46	Sequence 46, Appl
42	95	3.9	917	10 US-09-815-242-5603	Sequence 5603, Ap
43	95	3.9	920	10 US-09-815-242-12181	Sequence 12181, A
44	95	3.9	920	10 US-09-815-242-12995	Sequence 12995, A
45	95	3.9	920	10 US-09-815-242-13148	Sequence 13148, A

#### ALIGNMENTS

RESULT 1  
US-09-815-242-10253  
; Sequence 10253, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlssen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA, 011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10253  
; LENGTH: 453  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
; US-09-815-242-10253  
Query Match 30.5%; Score 735; DB 10; Length 453;  
Best Local Similarity 36.4%; Pred. No. 1.4e-57;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

```
QY 20 IHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDYVYAKMGIDEIVYSRONDGSIAPLN 79
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 24 IYDPPDIYOLELIERIFGRCHLFLAHESQIPKPGDFPNYIMKEDAVVYRKODGSIKALN 83
QY 80 VCRHGRKTLVSVENAGNAKGFVCSYHGWFGSGNGELQSVPEFKLDYGESLNKKGLGKEVA 139
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 84 QCHRRARVSVADCGNRAFCPYHGWSYGINGELIDPLPRAYPOGLCSHMGLEVP 143
QY 140 RVSEHGFITGCFQDEAPPLMDYLGDAWYLEPMF-KHSGLELVGPRGVITANKAP 198
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 144 CVSSYKGLIFGNMDSAPGLNDYLDGLAMYLDRREGGTEIVGYQKAVINCNMKEP 203
QY 199 AENFVGDVAVGVMTASSL-----RSGESIFSSLAGNAALP----- 234
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 204 AEGFASQYHALFSAHAAYVLGAKDGDSDKRLD-----GQIARVWETAKDALDFG 256
QY 235 --PEGAGLQMTSKYSGMGVLMGYSGVHSADLVPELMAFGAKOERLNKEIGDVARIV 292
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 257 QDGHSGGFETTERKDANVWV--DGAVSSYRETYAE-----AEQRLGEVRLRL 303
QY 293 RSLHNCVFPNNMSLJCSGVFKVNNPIDANTTEVWTYAIYVCKMPEDLKRLADSVQRTA 352
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 304 AGHNN--IFPTLSMLNGATLRLVHPPGPDQVEWAFCTIDKASDEYKKAEPNSATPAF 361
QY 353 GPAGFESDDNDNMETASQNGKKYQSRSDLSNLGFGEDYGDVAVYPGVYKSAIGETS 412
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 362 GPAGFLEQDDSENNCELQKLLKGRANRSKLCLEMGLQEKRRDGIPIGT-NYIFSETA 420
QY 413 YRGFYRAYQAHVSSSNWAEFEHASSTWHTLTK 445
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 421 ARGMYQRMADLLSSEWQEVLDKTAAYQOEYMK 453
```

## RESULT 2

```
US-09-815-242-11692
; Sequence 11692, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11692
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
```

US-09-815-242-11692

Query Match 15.6%; Score 375.5; DB 10; Length 385;  
Best Local Similarity 31.3%; Pred. No. 1e-25;  
Matches 105; Conservative 56; Mismatches 145; Indels 29; Gaps 14;

```
QY 34 IFARNMLFLTHDSLIPAPGDYVYAKMGIDEIVYSRONDGSIAPLWVCRHGRKTLVSV 93
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 3 IFGNWVFLAHESQIPQPGDYVITLTLGROPVITLRDKNELHALINSCAHRGAMLCRRKT 62
QY 94 GNAGFVCSYHGWFGSGNGELQSVPEEK-DLYGSL-NKKCLGLKEVARVESHGFIYGC 151
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 63 GNKNSFTCPFRHGWTFSSNKKLLKAKDDESTAYAPETFEHESHDQKIPRQSYRGFLFS 122
QY 152 FDOAPPLMDYLGDAWYLEPMFKHS-GGLELVGPPGVYIKANKKAPAEFNGDAYHVG 210
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 123 LNAVQPLEAVLGFTCKITIDLVDQAPEGLEVLGSSSYIEGNNWKLGAEN-GADGYHS 181
QY 211 --WTHASSL--RSGESIFSSLAGNAALPPEGAGLQMTSKY--SGMGVLMGYSGVHSA 263
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 182 VVHNNVASTMSRRMYEA-----EGHTVDANGWSKSLGGGCGFDNGHMLM-----TR 229
QY 264 DLVPELMAFGAKOERLNKEIGDVAR-RYRSLHNCVFPNNMSL--TCSGVFVWNPIDA 321
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 230 ALNPEVPRV-YAHRERLQAEFGERRADQMVETRNLCILYVNLMDQFSTQIRVIRPIAV 288
QY 322 NTTEV--WTYAIY-EKDMPEDLKRLADSVQRTAG 353
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 289 DKTEVITWCFAPKGSQDARALRIROYEDFPNVSG 323
```

## RESULT 3

```
US-09-738-626-6140
; Sequence 6140, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIRO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, MAKOTO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6140
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
```

Query Match 15.6%; Score 375; DB 9; Length 490;  
Best Local Similarity 26.3%; Pred. No. 1.6e-25;  
Matches 118; Conservative 66; Mismatches 215; Indels 50; Gaps 15;

```
QY 20 IHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDYVYAKMGIDEIVYSQNGSITAPLN 79
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 27 IFTDPELFELEMRHIFEGNWIYLAHESQIPNAGDYFTTYIGROPITWTSKSGDTLCLIN 86
QY 80 VCRHGRKTLVSVENAGNAKGFVCSYHGWFGSGNGELQSVPEEKD-LYGESLNKK-CIGLKE 137
```

```

Db      87  ACSHRAMLCRRKTDNRLLTLCPEHGWTFESNDGALLKVKDEKGAYPENFRDGSIDLRR 146
Qy      138 VAASEFHGTYGCGFOEAPPLMDYIGDAWYIEPMEFKHS--GGLVYGPBGKVIKANKK 196
Db      147 VPFEEISRGFLSLNDVYSLEEHGDRTRYIDMLVDSPGELFVLRSSSTIYTDGNKK 206
Qy      197 APAENFVGAYHGWTH-----ASSLRGSESIFFSLAGNALPPEGAGLQMTSKYSGSM 250
Db      207 LOLEN-GAGCYHVSSTHMYAATTSRGTGESNETKADAGCKWGGGCTYSPTGHML 265
Qy      251 GVLMDYSGVHASADLYPELMAFGAKOERLNKEIGDVRAR-IYRSLNCTVPPNNSML-T 308
Db      266 --LMMWGNPEDEPLF-----ERRDEFKKKEFGEGEFMYGASRNLCCLPNNYIMDQ 315
Qy      309 CSGVEFVNPIDANTTEWYTAIVEKMDPEDKRLADSVORTAGAGWESDNDNMET 368
Db      316 FSSQIRHPRISVDQTEVITYCTAPRGESAEARANRIROYEDFNATGMATPDDLEFRS 375
Qy      369 ASONGKKYOSRD--SDLSNLGFGEDVYGDVYPCGVGK-----SAIGETSYRGFY 417
Db      376 CQ---KTYGASAFPMWDM--TRGLGHQYQG---PREVAKGLGMNEVLSSGARTEDGLX 426
Qy      418 -----RAYQAHVSSSNWAEFEHAST 438
Db      427 PLOGHFWMHELMQEAANKOSIKERELADDT 455

```

## RESULT 4

```

US-09-815-242-5097
; Sequence 5097, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5097
; LENGTH: 424
; TYPE: prt
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5097

```

```

Query Match      14.0%: Score 336.5; DB 10; Length 424;
Best Local Similarity 27.3%: Pred No. 3.5e-22;
Matches 114; Conservative 70; Mismatches 178; Indels 55; Gaps 16;
Qy      25 ELFOHELKTIFARNMLFLTHDSLIPAPGDYVTAKMIDEIVSRONDGSIRAFANVCRRH 84

```

```

Db      24 ELHRRLEHIFDDSMLYAAHLSELREPGDFITRDYGGRNLIQRRADGEPAYVLNACAR 83
Qy      85 GKTIVSVEGNANKKGYCSTHGWGSGNGLQSVPEPKDIYGSLNKKC---LGLKEVARV 141
Db      84 GAKVCAERGNQSORFCPPHGWTHDSHSLIGLP-DKAAVQHA--GQCHPELSLTQVKUA 140
Qy      142 ESFPHGTYGCGFOEAPPLMDYIGDAWYIEPMEFKHS--GGLVYGPBGKVIKANKAPAE 200
Db      141 -YRNFLFTHYARQASLTETTYGAKDYIDLICDQSEALELTIIPGFEHSIRANKMLAE 199
Qy      201 NEVGDAHYHGWTHASSLSRGSSEIFSSLAGNALPPEGAGLQMTSKYSGSMGYLMDYSGV 260
Db      200 NGV-DAYHLFPFAHRYLEYLNTL-----GTPESHKRRGRG-EALGNHALI 244
Qy      261 HS-----ADLYVE-LMARGCAKOEERLNKEIGDVRAR-IYRSLNCTVPPN-- 303
Db      245 ISGPPSTGRPIAYWSPLEFPALEKPSIAKFERLVERFGGARAEDIAHTKSLFIPEPLV 304
Qy      304 NSMLTCSGVFKYWNPIDANTTEWYTAIVEKMDPEDKRLADSVORTAGAGWESDND 363
Db      305 NDLGLN--IRSFPTADEVSYTWGAGFADETREBARINGLISFTGPGGFGTDPDV 362
Qy      364 DNMEFASONGKKYOSRSDLSNLGFGEDVYGDVYPCGVGKSAIGETSYRGFYRAY 420
Db      363 ELLESCQ---RAYAH-----AALGYSDFSRG---MGPATRRHVEDEQNRGFWREM 406

```

## RESULT 5

```

US-09-776-490-19
; Sequence 19, Application US/09776490
; Patent No. US20010012886A1
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESS: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. US20010012886A1th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/776,490
; FILING DATE: 02-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/810,009
; FILING DATE: 04-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Spullin, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELELEX: 575102
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

SEQUENCE DESCRIPTION: SEQ ID NO: 19;  
US-09-776-490-19

Query Match 7.7%; Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%; Pred. No. 2.9e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGNSGE 113  
|||||  
DB 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGNSCK 35

RESULT 6  
US-09-776-491-19  
; Sequence 19, Application US/09776491  
; Patent No. US20010013135A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; Gray, John  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009  
; CITY: Charlotte  
; STATE: No. US20010013135A1th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/776.491  
; FILING DATE: 02-Feb-2001  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/810.009  
; FILING DATE: 04-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spruill, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: 5718-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; TELEEX: 575102  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-491-19

Query Match 7.7%; Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%; Pred. No. 2.9e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGNSGE 113  
|||||  
DB 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGNSCK 35

RESULT 7  
US-09-776-490-20  
; Sequence 20, Application US/09776490  
; Patent No. US20010012886A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776.490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810.009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-490-20

Query Match 7.2%; Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%; Pred. No. 2.7e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGNSGE 113  
|||||  
DB 1 NVCRHRGKTLVSEAGNAKGFVCSYHGWFSGNSCK 35

RESULT 8  
US-09-776-491-20  
; Sequence 20, Application US/09776491  
; Patent No. US20010013135A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; Gray, John  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009  
; CITY: Charlotte  
; STATE: No. US20010013135A1th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-491-20

Query Match 7.2% Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%; Pred. No. 2.7e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCNRHGKTLVSEAGNAGFCVSTHGMGFGSNGE 113  
|||||||:|||||||:|||||||:|||||||:  
DB 1 NVCNRHGKTLVNAEAGNAGPCVGHGMGFGSNGK 35

RESULT 9  
US-09-776-490-21  
Sequence 21, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-490-21

Query Match 7.0% Score 168; DB 10; Length 35;  
Best Local Similarity 80.0%; Pred. No. 9.4e-09;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCNRHGKTLVSEAGNAGFCVSTHGMGFGSNGE 113  
|||||||:|||||||:|||||||:|||||||:  
DB 1 NVCNRHGKTLVDAEAGNAGPCVGHGMGFGSNGK 35

RESULT 10  
US-09-776-491-21  
Sequence 21, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-491-21

Query Match 7.0% Score 168; DB 10; Length 35;  
Best Local Similarity 80.0%; Pred. No. 9.4e-09;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;



CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-491-12

Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0013;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAKGVCSTHGWFSGSNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NCCRHRGMRICRADGNAKAPTCSTHGWAYDTSAG 34

RESULT 14  
US-09-776-490-14  
Sequence 14, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-776-490-14

Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0016;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAKGVCSTHGWFSGSNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NCCRHRGMRICRADGNAKAPTCSTHGWAYDSAG 34

RESULT 15  
US-09-776-491-14  
Sequence 14, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-776-491-14

Query Match 4.5%: Score 109; DB 10; length 35;

Best Local Similarity 52.9%; Pred. No. 0.0016;

Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

Oy 79 NCRHRRKTLVSEAGNAKGFVCSYHGNGFGSNG 112

Db 1 NCRHRRGMRICRADGNGNAKSPTCSYHGWAYDSAG 34

Search completed: January 27, 2003, 09:20:55  
Job time : 7 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 4, 2003, 10:53:49 ; Search time 0.001 Seconds  
(without alignments)

455.328 Million cell updates/sec

Title: us-09-720-451-6

Perfect score: 2187

Sequence: 1 ARVITQLEFYDIDIFLKFVN.....VSLVNDYINGDLFAVFADF 408

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.1  
Ygapop 10.0 , Ygapext 0.1  
Fgapop 6.0 , Fgapext 0.1  
Delop 6.0 , Delext 0.1

Searched: 1 segs, 558 residues

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1 summaries

Command line parameters:  
-MODE=frame-p2n.model -DEV=soft -Q=us-09-720-451-6 -DB=us-09-720-451-5  
-SUFFIX=plc -OUT=align6\_5 -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=BITS  
-START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=1 -DOCALIGN=200  
-THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=plc  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO\_XLPHYX  
-NEG\_SCORES=0 -LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.1 -Fgapop=6  
-Fgapext=0.1 -Ygapop=10 -Ygapext=0.1 -Delop=6 -Delext=0.1

Database : us-09-720-451-5.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	998	45.6	558	1	us-09-720-451-5

#### ALIGNMENTS

RESULT 1  
us-09-720-451-5

Alignment Scores:

Pred. No.: 0  
Score: 998.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 45.63%  
DB: 1  
Length: 558  
Matches: 185  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

us-09-720-451-6 (1-408) x us-09-720-451-5 (1-558)

QY 4 IIEVALThGlnIeuPheTyrAspThrAspIlePheLeuLysPheValAsnAspCysArg 23  
DB 3 ATTGTCACCAATATTATTATGATACGATATATTCCTCAAAATTGTGTAACACGCTGTCCG 62

QY 24 GlnIleGlyIleThrCysProIleValProGlyIleMetProIleAsnAspTyrIleGly 43  
DB 63 CAAATTGGATAACAGTGTCTATTGCTACCTGCAATATATGCCATTATATTAACAAGGCG 122  
QY 44 PheIleArgMetThrGlyPheCysLysThrLysIleProIleAspIleMetIleAlaIleu 63  
DB 123 TTATATCCGATGACTGGGTTTGTGCAACAAACAAATACCAAGCTGACATTTATGGCTGCTTA 182  
QY 64 GluProIleLysAspAsnGluAlaIleValLysAlaTyrGlyIleHisLeuGlyThrGlu 83  
DB 183 GAGCCTATCAAGACATGAGACATGAGACGCTGCAAGGCTTATGAGATTCACCTGGCACTGAA 242  
QY 84 MetCysLysLysIleLeuAlaHisGlyIleLysThrLeuHisLeuTyrThrLeuAsnMet 103  
DB 243 ATGTGCAAAAGATTATTACCTCATGGAATTAACATTCATCTTATACACTAAATATAG 302  
QY 104 GluLysSerAlaLeuAlaIleLeuMetAsnLeuGlyLeuIleGluLysSerLysValSer 123  
DB 303 GAGAAATCTGCATTGGCAATACTAATGAACCTTGGCTTAATGAAGATGCCAAAGTTTCT 362  
QY 124 ArgSerLeuProThrParArgProAlaAsnValPheArgValLysGluAspValArgPro 143  
DB 363 AGGTCTTACCTTGAGAGAGCGCCCTGCCAAATGTTTCCGTTPAAGAGAAGATGCCGTCCA 422  
QY 144 IlePheTrpAlaAsnArgProLysSerTyrIleSerArgThrIleGlyTrpAspGlnTyr 163  
DB 423 ATCTTTTGGCAAAATGACCAAAAGCTCATATCAAGACCATAGGATGGGATCAATAC 482  
QY 164 ProHisGlyArgTrpLysAspSerCysAsnProSerTyrGlyAlaLeuSerAspTyrGln 183  
DB 483 CCACATGGGCGTTGGGCTGATTCCTGTAATCCATCATATGTCATTTATCTGATTATCAG 542  
QY 184 PheMetArgProArg 188  
DB 543 TTCATCGGCGCACGT 557

Search completed: March 4, 2003, 10:53:50  
Job time : 1 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:56:07 ; Search time 10.1667 Seconds  
(without alignments)  
1299.432 Million cell updates/sec

Title: US-09-843-250-33

Perfect score: 2408  
Sequence: 1 MNYNKILVSESGLSQKHLI.....AEFHASSTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6A.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/6ackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	224	9.3	446	US-09-004-393B-4	Sequence 4, Appl
2	223	9.3	439	US-09-004-393B-2	Sequence 2, Appl
3	185	7.7	35	US-08-810-009-19	Sequence 19, Appl
4	174	7.2	35	US-08-810-009-20	Sequence 20, Appl
5	168	7.0	35	US-08-810-009-21	Sequence 21, Appl
6	113.5	4.7	379	US-09-028-934-36	Sequence 36, Appl
7	110	4.6	35	US-08-810-009-12	Sequence 12, Appl
8	109	4.5	35	US-08-810-009-14	Sequence 14, Appl
9	108	4.5	35	US-08-810-009-13	Sequence 13, Appl
10	108	4.5	35	US-08-810-009-15	Sequence 15, Appl
11	108	4.5	622	US-09-311-626A-4	Sequence 8, Appl
12	106.5	4.4	432	US-08-809-326A-16	Sequence 16, Appl
13	106.5	4.4	649	US-08-809-326A-15	Sequence 15, Appl
14	106	4.4	35	US-08-810-009-9	Sequence 9, Appl
15	104	4.3	35	US-08-810-009-18	Sequence 18, Appl
16	103	4.3	35	US-08-810-009-11	Sequence 11, Appl
17	102	4.2	35	US-08-810-009-8	Sequence 8, Appl
18	102	4.2	35	US-08-810-009-10	Sequence 10, Appl
19	101	4.2	17	US-08-810-009-44	Sequence 44, Appl
20	101	4.2	17	US-08-810-009-45	Sequence 45, Appl
21	100	4.2	35	US-08-810-009-17	Sequence 17, Appl
22	99.5	4.1	256	US-09-325-932A-57	Sequence 57, Appl
23	98	4.1	5588	US-09-036-987A-6	Sequence 6, Appl
24	98	4.1	5588	US-09-370-700-6	Sequence 4800, Ap
25	96	4.0	563	US-09-134-001C-4800	Sequence 46, Appl
26	95	3.9	17	US-08-810-009-46	Sequence 1, Appl
27	94.5	3.9	350	US-08-828-922-1	Sequence 1, Appl

28	94	3.9	560	US-08-814-052-6	Sequence 6, Appl
29	94	3.9	560	US-08-812-829-6	Sequence 6, Appl
30	94	3.9	1693	US-08-478-507-7	Sequence 7, Appl
31	94	3.9	1693	US-09-128-275A-7	Sequence 7, Appl
32	94	3.9	1693	US-09-553-427-7	Sequence 7, Appl
33	93	3.9	35	US-08-810-009-16	Sequence 16, Appl
34	93	3.9	421	US-09-239-303-2	Sequence 12, Appl
35	93	3.9	631	US-09-345-468-12	Sequence 12, Appl
36	93	3.9	525	US-09-414-453A-12	Sequence 12, Appl
37	92.5	3.8	525	US-08-348-891A-2	Sequence 2, Appl
38	92.5	3.8	525	US-08-805-817-2	Sequence 2, Appl
39	92	3.8	1087	US-08-570-311-8	Sequence 8, Appl
40	92	3.8	1087	US-08-353-485-8	Sequence 8, Appl
41	92	3.8	1358	US-08-570-311-27	Sequence 27, Appl
42	92	3.8	1693	US-08-840-316-1	Sequence 1, Appl
43	92	3.8	1693	US-08-809-523-1	Sequence 1, Appl
44	92	3.8	1693	US-08-471-971-1	Sequence 1, Appl
45	92	3.8	1693	US-09-402-776-1	Sequence 1, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-004-393B-4
; Sequence 4, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
; APPLICANT: Rathnasabapathi, Balu
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; TITLE OF INVENTION: Plants Transformed Therewith
; FILE REFERENCE: UF-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-004-393B-4

Query Match          9.3%; Score 224; DB 4; Length 446;
Best Local Similarity 27.4%; Pred. No. 56-13;
Matches 58; Conservative 37; Mismatches 97; Indels 20; Gaps 6;

QY 11 ESGLSQKHLIHGDEELFOHELKTIFARNMLFLHDSLIAPAGDYVAKMGIDEIVISRON 70
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 99 EDALTPPSWTYTPAPAFYSHELRIFYKGMGVAGYSQVKEKNQYFTGSGANVEYLVSRG 158
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 71 DGSIRAFIYVCRHGRKTVLSEVAGNAKGFVCSYHGMFGSNGELQSVPEKIDYGSLNK 130
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 159 QGELHAFHNVCTHRA-SILAQSGKSKSCFYCRHGVYGLDGLAKA--SKARETQNDP 215
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 KCLGLKEVARVESFHCFTYGCQDEAPPLMD---YLQDA---AWLEPMFKHSGGLE 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 KETGLAPL-KVAVMGFFIILSDRSIDANADVETEWIGKSAEDVKAHARDPNIKFTHRSE 274
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 LVGPPGKVIYKAMWKAFAENFVGDAYHVGWTH 213
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 F-----PMECNMKVFCNDYIDSSYHVPYAH 299
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-09-004-393B-2
; Sequence 2, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
```

```

? GENERAL INFORMATION:
? APPLICANT: Briggs, Steven P.
? APPLICANT: Johnal, Gurmukh S.
? APPLICANT: Gray, John
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
? TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
? NUMBER OF SEQUENCES: 65
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: BELL, SELTZER, PARK & GIBSON
? STREET: P.O. Drawer 34009
? CITY: Charlotte
? STATE: No. 6211437th Carolina
? COUNTRY: USA
? ZIP: 28234
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/810,009
? FILING DATE: 04-MAR-1997
? CLASSIFICATION: 800
? ATTORNEY/AGENT INFORMATION:
? NAME: Spruill, W. Murray
? REGISTRATION NUMBER: 32,943
? REFERENCE/DOCKET NUMBER: 5718-4
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-881-3140
? TELEFAX: 919-881-3175
? TELEX: 575102
?
? INFORMATION FOR SEQ ID NO: 20:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 35 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
? US-08-810-009-20
?
? Query Match 7.2%; Score 174; DB 4; Length 35;
? Best Local Similarity 85.7%; Pred. No. 6 1e-10;
? Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0
?
? QY 79 NVCNRHKGKTLVSYEAGNAGKGFVCSYHGWGFGSNGE 113
? ||||| 111111 || 111111111111
? Db 1 NVCNRHKGKTLVNAEAGNAGKGPVCGYHGWGFGSNGK 35

```



```
; Sequence 21, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Gray, John
; APPLICANT: Jomal, Gurmukh S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ. ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-21

Query Match          7.0%; Score 168; DB 4; Length 35;
Best local similarity 80.0%; Pred. No. 2.3e-09;
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 MVCRRHGKTLVSEAGNAGKGVCSYHGMGSGNGE 113
DB 1 MVCRRHGKTLVSEAGNAGKGVCSYHGMGSGNGK 35

RESULT 6
US-09-028-934-36
; Sequence 36, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrolytic Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ. ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-028-934-36

Query Match          4.7%; Score 113.5; DB 3; Length 379;
Best local similarity 20.1%; Pred. No. 0.014;
Matches 96; Conservative 50; Mismatches 144; Indels 187; Gaps 26;

QY 20 IHGDELFQ----HEKTIIPARN--LFLTHSLIPAGDGYTAKGIDELVYSRONDS 73
DB 5 IHQEPRIQASGVNDLTQTASWYVAMRSDALRGKP--VAIKLFCQPLVAMRDGGR 61

QY 74 IRAFVNCRRHCKTLVSEAGNAGKGFV-----CSYHGMGFGSNGSLQSVFPFKDLYES 127
DB 62 PYVMERYCSHLGASL-----AKGVVEGCIQCPRHNMRYDSTGACSHNP-----GHS 108

QY 128 LNKCKGLKEVARVESF-----HGFIYGCDFQEAR-----PLMDYLGDAMV 169
DB 109 -----TEVRLPEIPTARQSYVPMERKGFVWVGTRAPLPLEMPAESSESH 160

QY 170 LEPMEKHSGLGLVPPGKAVIKANKKAPAEFVGDAYHYGWTTHASSLSRGE-SIF---S 225
DB 161 QSLRFAYETTTSVL-----RII-----ENFY-DAQHAAPVHQPLISAFELKLFDESS 206

QY 226 SLAGNALRPE-----GAGIOM-TSKYSGMGVL-----MNGSYGVHSADLV 266
DB 207 PPPGQALARDGAWFGAGIDFHYDRYEGPLGYISRTGLGLSMSRMQLHFDGYPG----- 259

QY 267 PELMAFGAKOERLNKEIDVRAIRYRSHLNCYVFPNNSMLTCSGVFKWNPIDANTTREV 326
DB 260 -----GCIMTYSLD---GDVKYRL---LQCY-----TPDKRETYWA 289

QY 327 WTYAIVKEMPEDLKRRLADSVQRTA-----GPAF---WESDDNDNMETAQ 371
DB 290 HMLLAIRK-----DDGVRSANFLIYGLQTMWAAAGYDVIAIWMKAD----- 332

QY 372 NKKKYRSRDLNSNGFEDYGDVAVYGVYCKSAIGETSYRGFRARQAAHYSSSN 428
DB 333 GGGAFSKYDOLL-----KYRAFRRVNVKVALEN 362

RESULT 7
US-08-810-009-12
; Sequence 12, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
```

```

; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-810-009-12
;
Query Match          4.6%; Score 110; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.00079;
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRHGRKTLVSEAGNAKGFVCSYHGWFSGSNG 112
      1 11111 : :11111 11111 : : 1
      1 NCRHRGMRICRADAGNAKAPTCSYHGWAYDTAG 34

Db
;
; RESULT 8
; US-08-810-009-14
; Sequence 14, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
```

```

; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
;
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-810-009-14
;
Query Match          4.5%; Score 109; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.00098;
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHGRKTLVSEAGNAKGFVCSYHGWFSGSNG 112
      1 11111 : :1111 11111 : : 1
      1 NCRHRGMRICRADGNAKSPCTCSYHGWAYDSAG 34

Db
;
; RESULT 9
; US-08-810-009-13
; Sequence 13, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
;
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-810-009-13
;
Query Match          4.5%; Score 108; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.0012;
```



```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,326A
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 224711/94
; FILING DATE: 20-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106006/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106008/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106009/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106010/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106011/95
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Miller, Charles E.
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 7426-043-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-809-326A-16

Query Match 4.4%; Score 106.5; DB 4; Length 432;
Best Local Similarity 21.4%; Pred. No. 0.079;
Matches 53; Conservative 39; Mismatches 89; Indels 67; Gaps 11;

QY 195 WKAPAEENVGDYVHGWTHASSL-----RSGESIFSSLAGNAALPPEGAGLOMTSKY 246
DB 22 WNLPAD-----LAFKRTLTLPKPYIMGRHTWESIGRPLPGR-----KNITLSSQP 66

QY 247 GSGMGVLMDSYGVHSADL-----VPELMAFGAKOERLNKEIGDVRARIYRSHLCT 299
DB 67 GTDRVYVW-----VKSVDIAIACGDVEIMVIGG---RYEQLPRAKOKLYLTHIDAE 118

QY 300 V-----FPNNSMLTCSGYEFKAWNPIDANTTWTYATYEKMPEDLKRRLADSVORTGCP 354
DB 119 VEGDTHRPDVEPPDDWESVFSEFHDAQONSHSYEFLE-----RRILSISSSSGP 170

QY 355 AGFWESDDNDNME-----TASONGKKKYQSRDSD-----LSNLGCGEYVDDAVYGPVVG 404
DB 171 -----DNQKNIMSOVLSTPGVPPQDKLSGNETKQIQOTRQGNTEMESDATIAGASG 224

QY 405 KSAIGETS 412
DB 225 KDKTSSST 232
```

```

; APPLICANT: Obara, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,326A
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 224711/94
; FILING DATE: 20-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106006/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106008/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106009/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106010/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106011/95
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Miller, Charles E.
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 7426-043-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-809-326A-15

Query Match 4.4%; Score 106.5; DB 4; Length 649;
Best Local Similarity 21.4%; Pred. No. 0.15;
Matches 53; Conservative 39; Mismatches 89; Indels 67; Gaps 11;

QY 195 WKAPAEENVGDYVHGWTHASSL-----RSGESIFSSLAGNAALPPEGAGLOMTSKY 246
DB 22 WNLPAD-----LAFKRTLTLPKPYIMGRHTWESIGRPLPGR-----KNITLSSQP 66

QY 247 GSGMGVLMDSYGVHSADL-----VPELMAFGAKOERLNKEIGDVRARIYRSHLCT 299
```

```

Db      67 GTDDRVTW-----VKSVDIAIACGDPPEIMVIGG---RVYEQPLPKAQKLYLTHIDAE 118
QY      300 V-----FPNNSMLTCSGVFEVWNPNDANTTEVTVTAIVEDMEDJRRRLADSVQTAGP 354
      119 VESDTHPPDEPPDMSVSEFHDADQNSHYEFITL-----RILMSISSSGP 170
QY      355 AGFWSDDNDNME-----TASQNGKKYOSRPSD-----LLSNLGFEDVYGDVYGVYG 404
      171 -----DNQKNINSQVLTSPQGVPPQDKLSGNETKQIQOTRQGNKTNEMESDATINGASG 224
Db      405 KSAIGETS 412
      225 KDKTSSPT 232

RESULT 14
US-08-810-009-9
; Sequence 9, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-9

Query Match      4.4%; Score 106; DB 4; Length 35;
Best Local Similarity 51.4%; Pred. No. 0.0019;
Matches 18; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY      79 NVCRHRGKTLVSEAGNAGKGFVCSYHGSGSNGE 113
      1 NCRHRGMRICRSDAGNAKAPTCSTYHGWAYDIAGK 35

RESULT 15
US-08-810-009-18
; Sequence 18, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:

```

```

; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-18

Query Match      4.3%; Score 104; DB 4; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.003;
Matches 18; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY      79 NVCRHRGKTLVSEAGNAGKGFVCSYHGSGSNG 112
      1 NSCRHRGALLCPSPKGNQKFWCRHGWSTYDSG 34

```

Search completed: January 27, 2003, 09:04:54  
Job time : 11.1667 secs

10

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:54:32 ; Search time 25.8333 Seconds  
(without alignments)  
3581.232 Million cell updates/sec

Title: US-09-843-250-33  
Perfect score: 2408  
Sequence: 1 MNYNKILVSESGLSOKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21:\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.protozoan:\*  
12: sp.virus:\*  
13: sp.yeast:\*  
14: sp.unclassified:\*  
15: sp.rvirophage:\*  
16: sp.bacteriophage:\*  
17: sp.archaeovirus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2330	96.8	449	2	Q938R9
2	2290	95.1	449	2	Q52142
3	2206	91.6	449	2	Q92173
4	2191	91.0	447	2	Q915D2
5	2187	90.8	447	2	Q52382
6	2071	86.0	447	2	Q8VUP4
7	2051	85.2	447	2	P95564
8	2022	84.0	447	2	Q8RT14
9	1968.5	81.7	451	2	Q45695
10	1490	61.9	277	2	Q9ETK2
11	1488	61.8	277	2	Q9F5S4
12	1486	61.7	277	2	Q9F5T3
13	1484	61.6	277	2	Q9F5S8
14	1483	61.6	277	2	Q9F5S9
15	1451	60.3	277	2	Q9F5S3
16	1405	58.3	277	2	Q9EUC7

17	1403	58.3	277	2	Q9F5T8	Q9F5T8 pseudomonas
18	1402	58.2	277	2	Q9F5T5	Q9F5T5 pseudomonas
19	1399	58.1	277	2	Q9ETV3	Q9ETV3 pseudomonas
20	1399	58.1	277	2	Q9F5S5	Q9F5S5 pseudomonas
21	1396	58.0	277	2	Q9F5T6	Q9F5T6 pseudomonas
22	1396	58.0	277	2	Q9F5S7	Q9F5S7 pseudomonas
23	1393	57.8	277	2	Q9F5T4	Q9F5T4 pseudomonas
24	1393	57.8	277	2	Q9F5T2	Q9F5T2 pseudomonas
25	1393	57.8	277	2	Q9F5T1	Q9F5T1 pseudomonas
26	1393	57.8	277	2	Q9F5T0	Q9F5T0 pseudomonas
27	1393	57.8	277	2	Q9F5S6	Q9F5S6 pseudomonas
28	1390	57.7	277	2	Q9F5T7	Q9F5T7 pseudomonas
29	1384	57.5	277	2	Q9F5S2	Q9F5S2 pseudomonas
30	1353.5	56.2	450	2	Q92H33	Q92H33 burkholderi
31	1139	47.3	226	2	Q8VUM6	Q8VUM6 ralstonia s
32	1134	47.1	214	2	Q8VR23	Q8VR23 pseudomonas
33	1134	47.1	214	2	Q8VL21	Q8VL21 pseudomonas
34	1127	46.8	214	2	Q8VR24	Q8VR24 pseudomonas
35	1123	46.6	214	2	Q8VR22	Q8VR22 pseudomonas
36	1111.5	46.2	455	2	Q93NA8	Q93NA8 burkholderi
37	1067	44.3	214	2	Q8VR25	Q8VR25 burkholderi
38	1043.5	43.3	437	2	Q9WXG8	Q9WXG8 alcaligenes
39	1040	43.2	208	2	Q9F6B6	Q9F6B6 marinobacte
40	1039	43.1	206	2	Q8VUM5	Q8VUM5 naphthalene
41	1038	43.1	214	2	Q8VR21	Q8VR21 burkholderi
42	1038	43.1	303	2	Q8VUM2	Q8VUM2 ralstonia s
43	1016	42.2	297	2	Q93M40	Q93M40 uncultured
44	1009	41.9	297	2	Q93M39	Q93M39 uncultured
45	1002	41.6	297	2	Q93M41	Q93M41 uncultured

ALIGNMENTS

RESULT 1	
Q938R9	PRELIMINARY: PRT: 449 AA.
ID Q938R9	
AC Q938R9:	
DT 01-DEC-2001 (TREMBLrel. 19, Created)	
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)	
DE Naphthalene dioxygenase.	
GN NAHAC.	
OS Pseudomonas fluorescens.	
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;	
OC Pseudomonas.	
OX NCBI_TaxID=294;	
RN [1]	
RP SEQUENCE FROM N.A.	
RL Min K -H, JI S -H;	
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL: AY048759; AAL07262.1; -	
DR InterPro: IPR001281; Rleske.	
DR InterPro: IPR001663; Ring_hydroxyl_A.	
DR Pfam: PF00355; Rleske; 1.	
DR Pfam: PF00848; Ring_hydroxyl_A; 1.	
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.	
DW Dioxigenase.	
SO SEQUENCE 449 AA; 49608 MW; 5EFDDE0282FE812 CRC64;	
Query Match	96.8%; Score 2330; DB 2; Length 449;
Best Local Similarity	96.0%; Pred. No. 6.1e-171;
Matches 431; Conservative 12; Mismatches 6; Indels 0; Gaps 0;	
QY 1 MNYNKILVSESGLSOKHLIHDELFQHELTIFARMMLFLTHDSLIPADGVYATKMG 60	
DB 1 MNYNKILVSESGLSOKHLIHDELFQHELTIFARMMLFLTHDSLIPADGVYATKMG 60	
QY 61 IDEVIVSNDGSIATFLNVCRRHCKTLVSVDAGAKGFVCSYHGWSGSGELOSYPE 120	
DB 61 IDEVIVSNDGSIATFLNVCRRHCKTLVSVDAGAKGFVCSYHGWSGSGELOSYPE 120	
QY 121 KDLVESLKKKCLGLKEVARVESFHGFIYGCEDGAPPLMDYLGDAANYLDEPMFRHSGSL 180	

```

|||||
121 KELYGESLNKKCLGKLEVARVESFHGFIYCGFQDEAPSLMDYLGDAAWYLEPIFKHSGGL 180
OY 181 ELVGPFGKVYIKAMWKAPEFVGDAYHVGWTHASSLSRSGSIFSSLAGNAALPPEGGL 240
Db 181 ELVGPFGKVYIKAMWKAPEFVGDAYHVGWTHASSLSRSGSIFSSLAGNAALPPEGGL 240
OY 241 QMTSKYSGMGVLMWDGYSGVHSAADLPVELMAFGAKOERLNKEIGDVRARIYRSHLNCV 300
Db 241 QMTSKYSGMGVLMWDGYSGVHSAADLPVELMAFGAKOERLNKEIGDVRARIYRSHLNCV 300
OY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTAGPAGFWS 360
Db 301 FPNNSVLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADAVQRTGPAAGFWS 360
OY 361 DDNDNMETAQNGKKYRSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
Db 361 DDNDNMETAQNGKKYRSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
OY 421 QAHVSSNMAEFEDASSTWHTELTKTTDR 449
Db 421 QAHVSSNMAEFEDASSTWHTELTKTTDR 449

```

RESULT 2  
052142 PRELIMINARY: PRT: 449 AA.

```

ID 052142 008194:
AC 052142: 008194: (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Iron-sulfur protein large subunit.
GN PAHAC.
OS Pseudomonas putida
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303:
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=OUS82;
RC Takizawa N.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RX MEDLINE=94209249; Pubmed=8157615;
RA TAKIZAWA N., Kaide N., Torigoe S., Moritani T., Sawada T., Satoh S.,
RA Kiyonara H.;
RT "Identification and characterization of genes encoding polycyclic
RT aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon
RT dihydrodiol dehydrogenase in Pseudomonas putida OUS82.";
RL J. Bacteriol. 176:2444-2449(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA Noboru T., Toshiya I., Takashi S., Kazuhiro Y., Masamichi K.,
RA Yue-Wu W., Masao F., Hohzoh K.;
RT "The molecular analysis of NAH7-type cluster located on the
RT chromosomes of Pseudomonas aeruginosa PAk1 and Pseudomonas putida
RT OUS82.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB004059; BAA20391.1;
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIHXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 449 AA; 49361 MW; 80020F54AAB1BEA CRC64;

```

Query Match 95.1%; Score 2290; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 7.3e-168;

Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

```

OY 1 MNYNKKILVSESGLSQKHLHGDEELFQHELTFFARNWLFTHDSLIAPAGDYVAKMG 60
Db 1 MNYNKKILVSESGLSQKHLHGDEELFQHELTFFARNWLFTHDSLIAPAGDYVAKMG 60
OY 61 IDEVIVSRONDGSIRAFELNVCRRHCKTLVNEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
Db 61 IDEVIVSRONDGSIRAFELNVCRRHCKTLVNEAGNAKGFVCSYHGMGFGSNGELQSVPE 120
OY 121 KDLYESLNKKCLGKLEVARVESFHGFIYCGFQDEAPSLMDYLGDAAWYLEPIFKHSGGL 180
Db 121 KDLYESLNKKCLGKLEVARVESFHGFIYCGFQDEAPSLMDYLGDAAWYLEPIFKHSGGL 180
OY 181 ELVGPFGKVYIKAMWKAPEFVGDAYHVGWTHASSLSRSGSIFSSLAGNAALPPEGGL 240
Db 181 ELVGPFGKVYIKAMWKAPEFVGDAYHVGWTHASSLSRSGSIFSSLAGNAALPPEGGL 240
OY 241 QMTSKYSGMGVLMWDGYSGVHSAADLPVELMAFGAKOERLNKEIGDVRARIYRSHLNCV 300
Db 241 QMTSKYSGMGVLMWDGYSGVHSAADLPVELMAFGAKOERLNKEIGDVRARIYRSHLNCV 300
OY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTAGPAGFWS 360
Db 301 FPNNSVLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADAVQRTGPAAGFWS 360
OY 361 DDNDNMETAQNGKKYRSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
Db 361 DDNDNMETAQNGKKYRSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
OY 421 QAHVSSNMAEFEDASSTWHTELTKTTDR 449
Db 421 QAHVSSNMAEFEDASSTWHTELTKTTDR 449

```

RESULT 3

```

ID 092173 PRELIMINARY: PRT: 449 AA.
AC 092173:
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase Fe-S large subunit.
GN NAHAC.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AN10;
RX MEDLINE=9365311; Pubmed=10433976;
RA Bosch R., Garcia-Valdes E., Moore E.R.B.;
RT "Genetic characterization and evolutionary implications of a
RT chromosomally encoded naphthalene-degradation upper pathway from
RT Pseudomonas stutzeri AN10.";
RL Gene 236:149-157(1999).
DR EMBL: AF039533; AB02136.1;
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIHXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KM Dioxygenase.
SQ SEQUENCE 449 AA; 49802 MW; BA510FD1B6f66E63 CRC64;

```

Query Match 91.6%; Score 2206; DB 2; Length 449;  
Best Local Similarity 89.5%; Pred. No. 2.1e-161;  
Matches 402; Conservative 25; Mismatches 22; Indels 0; Gaps 0;

OY 1 MNYNKKILVSESGLSQKHLHGDEELFQHELTFFARNWLFTHDSLIAPAGDYVAKMG 60



Db	1	MYNKNTLVTSSTGLTQXRLHLDGPEELFQRELETTIPARNMLFTIHDSLIFSPGDYITAKNG	60
Qy	61	IDEIVSRQNDGSIRAFELNVCRRHGKTLTVSEAGNAKGFCVSYHGMGFCFSGNGLQSVPE	120
Db	61	VDEIVSRQNDGSIRAFELNVCRRHGKTLTVHAEGNAKGFCVSYHGMGFCANGELQSVPE	120
Qy	121	KDLVGSLSLKNKCCGLGLEKVARVESHGFTTCYCFDQEPALPLMDYTGDAWYLEPFRKSGGL	180
Db	121	KELVGEALDKKCMGLKEVARVESHGFTTCYCFDEEAPSLDYMGDGAGWMEPFRKSGGL	180
Qy	181	ELVGPCKVYVTKAWKPAEKEFVGDAYHVGWMTAASSLRGSETFSSLAGMAALPREGAGL	240
Db	181	ELVGPCKVYVTKAWKPAEKEFVGDAYHVGWMTAASSLRTOQSVFTSLAGMAALPREGAGL	240
Qy	241	QMTSKYSGGMVLMDGYSGVHSAVLVELMAFGCAQOEKLEIGVRARIYRSHLNCYV	300
Db	241	QMTSKYSGGMVLMDGYSGVHSAVLVELMAFGCAQOEKLEIGVRARIYSHLNGYV	300
Qy	301	FPNNSMLTSCGVFVKNPPIDANTTEVWTYIVAEKMDPEDLKRLRLADSVORTAGPAFWES	360
Db	301	FPNNSFLTSCGVFVKNPPIDANTTEVWTYIVAMVEKMDPEDLKRLRLADSVORTFCPPAGFWES	360
Qy	361	DDDNMMETASONGKKYYSRSDLSLNLGFEEDYGDAAVYPGVVGKSAIGETSYRGFYRAY	420
Db	361	DDDNMMETESONAKKYYOSRSDGLSLNLGFEEDYGDDEVYPGVVGKSAIGETSYRGFYRAY	420
Qy	421	QAHVSSSNWAAEFEEHASTWHTLTKTTDR	449
Db	421	GAHLSSSSWAAEFEDASKNMHTLTKTTDR	449

RESULT 4			
ID	09L5D2	PRELIMINARY;	PRT; 447 AA.
AC	09L5D2;		
DT	01-OCT-2000 (TREMBLrel, 15, Created)		
DT	01-OCT-2000 (TREMBLrel, 15, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel, 15, Last annotation update)		
DE	polyaromatic hydrocarbon dioxygenase large subunit.		
GN	PAHAC.		
OS	Comamonas testosteroni (Pseudomonas testosteroni).		
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas		
OX	NCBI_TaxID=285;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN:H;		
RX	MEDLINE=21306905; PubMed=11414329;		
RA	Moser R., Stahl U.;		
RT	"Insights into the genetic diversity of initial dioxygenases from PAH		
RL	degrading bacteria.";		
RL	Appl. Microbiol. Biotechnol. 55:609-618(2001).		
DR	EMBL; AF252550; AAF72976.1; -;		
DR	HSSP; P23094; INDO.		
DR	InterPro; IPR001281; Rieseke.		
DR	InterPro; IPR001663; Ring_hydroxyl_A.		
DR	Pfam; PF00355; Rieseke; 1		
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.		
DR	PRINTS; PR00090; RINGDIOXGNASE.		
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.		
RW	Dioxygenase.		
QO	SEQUENCE 447 AA; 49548 MW; 2950EF36123A6F21 CRC64;		
Query Match	91.0%; Score 2191;	DR 2;	Length 447;
Best Local Similarity	88.9%;	Pred. No. 3e-160;	
Matches 399;	Conservative 27;	Mismatch 21;	Indels 2; Gaps
QY	1 MNYNNKILVSESGLSQKHLIHGDEELFOHSLKTIIFARNWELFTHDSLIPAPGDYATKMG 60		
Db	1 MIYEN-LVSSAGILQKHLIHGDEELFOHSLKTIIFARNWELFTHDSLIPSGDYATKMG 58		
QY	61 IDEYIVSRQNDGSTRFLNVCRIHKGTLVSEVAGNAKGFVCSYHGSGNGLQSVPE 120		

Db	59	VDEYIVSNQNGSVRAFLNVCNRHKGKTLVHAHAGNANGAFVCSYHGMFGSNGELQSVPE	118
Qy	121	KDLYGESLNKKCLGKLEKVARVESFHGELYGCEDQEAPLMDYIGDAAMYLEPMFKHSGL	180
Db	119	KEIYGDAIKKCKLGLKEVRIESEFHGELYGCEDQEAAPLIDYLGDAAWYLEPIFKHSGL	178
Qy	181	ELVGPCKVYVTKAMWKPAEKNFVDADVHVGTHASSLRSGSESTSSLAGNAALPREBAGL	240
Db	179	ELVGPCKVYVTKAMWKPAEKNFVDADVHVGTHAASSLRSGSESTTPLAGNAALPREBAGL	238
Qy	241	QMTSKYSGMGVLNDGYSGVISADLYPELMAFGAKOERLNKEIGDVARIRYRSHLCTV	300
Db	239	QMTSKYSGMGVLNDATSGHSADLYPELMAFGAKOEBKLKEIGDVARIRYRSHLCTV	298
Qy	301	FPNNSMLTCSGVFKWMPIDANTTEVMTYALVEKDMPEDLKRLADSVOPTAGPAGES	360
Db	299	FPNNSILTCSGVFKWMPIDENTEVMTYALVEKDMPEDLKRLADAVQRTGAPAGES	358
Qy	361	DDNDNMEIASONGKKYKORSDDLNTGFGEDVYGDVYFGVYKSAIGETSTRGFPAY	420
Db	359	DDNDNMEIETSONARKYOSNSDDLANTGFGEDVYGDVYCGVYKSAIGETSYRGFPAY	418
Qy	421	QAHVSSSNMAEEFHAASSTWHELTCTDR	449
Db	419	QAHVSSSNMAEEFENTSRWHELTCTDR	447

ID	052382	PRELIMINARY;	PRT;	447 AA.
AC	052382;			
DT	01-JUN-1998 (TREMBLrel, 06, Created)			
DT	01-JUN-1998 (TREMBLrel, 06, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel, 19, Last annotation update)			
DE	Naphthalene 1,2 dioxygenase large oxygenase component.			
CN	NAOAC.			
OS	Ralstonia sp. U2.			
OG	Plasmid pMW02.			
OC	Bacteria; Proteobacteria; beta subdivision; Ralstonia group;			
CC	Ralstonia.			
OX	NCBI_TaxID=70356;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=U2;			
RA	MEDLINE=98233751; PubMed=9573207;			
RT	Fuennayor S.L., Wild M., Boyes A.L., Williams P.A.;			
RT	"A gene cluster encoding steps in conversion of naphthalene to			
RL	gentisate in Pseudomonas sp. strain U2.";			
RN	J. Bacteriol. 180:2522-2530(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=U2;			
RA	MEDLINE=20576173; PubMed=11133965;			
RT	Zhou N.Y., Fuennayor S.L., Williams P.A.;			
RT	"nag genes of ralstonia (formerly pseudomonas) sp. strain U2 encoding			
RL	enzymes for gentisate catabolism.";			
RL	J. Bacteriol. 183:700-708(2001).			
DR	EMBL; AF036940; A012610.1; --.			
DR	HSSP; P23094; INDO.			
DR	InterPro; IPR001281; Rtske.			
DR	InterPro; IPR001663; Ring_hydroxyl_A.			
DR	Pfam; PF00355; Rtske; 1.			
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.			
DR	PRINTS; PR00090; RINGDIOXNAS.			
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.			
RW	Dioxygenase; Plasmid.			
SO	SEQUENCE 447 AA; 49570 MW; 4553AAP4B4410ED0 CRC64;			
Query Match	90.8%; Score 2187; DB 2; Length 447;			
Best Local Similarity	89.1%; Pred. No. 6,1e-160;			
Matches 400; Conservative 25; Mismatches 22; Indels 2; Gaps				

```
Db 1 MIYEN--LVSEAGLTQKHLIHGDELFQHELTIFARNWMLFTHDSLIPSPGDYVYAKMG 58
QY 61 IDEVIVSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELQSVPE 120
Db 59 VDEVIVSRQNDGSIIRAFILNVCRRHGRKTLVHAEGNAKGFVCSYHGWGSGNGELQSVPE 118
QY 121 KDLVGEISLNKKCLGKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 180
Db 119 KELYGDIKKKCLGKEVPRIESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 178
QY 181 ELVGPGRKVVIRKAMKKAPEAENFVGDAYHVGMTTHASSLSRGSSTFSSLAGNAALPEPEGAGL 240
Db 179 ELVGPGRKVVIRKAMKKAPEAENFVGDAYHVGMTTHASSLSRGSSTFSSLAGNAALPEPEGAGL 238
QY 241 QMTSKYSGGMVLDGYSGVHSAIDLVPMLAFGAKOEKRLKEIGDVARIRYSHLNTCTV 300
Db 239 QMTSKYSGGMVLDGYSGVHSAIDLVPMLAFGAKOEKRLKEIGDVARIRYSHLNTCTV 298
QY 301 FPNNSMLTSCGVFKWMPIDANTTEVMTYAIVEKDMPEDLKRLADSVQRTAGPAGFWES 360
Db 299 FPNNSMLTSCGVFKWMPIDANTTEVMTYAIVEKDMPEDLKRLADSVQRTAGPAGFWES 358
QY 361 DDNDNMETASONGKRYKOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
Db 359 DDNDNMETASONGKRYKOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 418
QY 421 QAHVSSSNMAEFHASSHTWHTELTKTTDR 449
Db 419 QAHVSSSNMAEFHASSHTWHTELTKTTDR 447
```

RESULT 6  
Q8YVUD4 PRELIMINARY; PRT; 447 AA.

```
AC Q8YVUD4;
DT 01-MAR-2002 (TremBrel. 20, Created)
DT 01-MAR-2002 (TremBrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBrel. 21, Last annotation update)
DE DntAC.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OX Burkholderia.
NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R34.
RX MEDLINE=20254695; PubMed=10795678;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Properties of the trihydroxytoluene oxygenase from Burkholderia
RT cepacia R34: an extradiol dioxygenase from the 2,4-dinitrotoluene
RT pathway.";
RL Arch. Microbiol. 173:86-90(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=R34;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Origins of the 2,4-dinitrotoluene pathway.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF169302; AA150021.1; -.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00335; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIPOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.
SQ SEQUENCE 447 AA; 49468 MW; 2A5BF8558320275C CRC64;
```

Query Match 86.0%; Score 2071; DB 2; Length 447;  
Best local similarity 83.5%; Pred. No. 5,1e-151;  
Matches 375; Conservative 36; Mismatches 34; Indels 2; Gaps 1;  
QY 1 MNYNKILVSESGLSQKHLIHGDELFQHELTIFARNWMLFTHDSLIPADGVYVYAKMG 60

```
Db 1 MSYQN--LVSEAGLTQKHLIHGDELFQHELTIFARNWMLFTHDSLIPSPGDYVYAKMG 58
QY 61 IDEVIVSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELQSVPE 120
Db 59 VDEVIVSRQNDGSIIRAFILNVCRRHGRKTLVHAEGNAKGFVCSYHGWGSGNGELQSVPE 118
QY 121 KDLVGEISLNKKCLGKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 180
Db 119 KELYGDIKKKCLGKEVPRIESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 178
QY 181 ELVGPGRKVVIRKAMKKAPEAENFVGDAYHVGMTTHASSLSRGSSTFSSLAGNAALPEPEGAGL 240
Db 179 ELVGPGRKVVIRKAMKKAPEAENFVGDAYHVGMTTHASSLSRGSSTFSSLAGNAALPEPEGAGL 238
QY 241 QMTSKYSGGMVLDGYSGVHSAIDLVPMLAFGAKOEKRLKEIGDVARIRYSHLNTCTV 300
Db 239 QMTSKYSGGMVLDGYSGVHSAIDLVPMLAFGAKOEKRLKEIGDVARIRYSHLNTCTV 298
QY 301 FPNNSMLTSCGVFKWMPIDANTTEVMTYAIVEKDMPEDLKRLADSVQRTAGPAGFWES 360
Db 299 FPNNSMLTSCGVFKWMPIDANTTEVMTYAIVEKDMPEDLKRLADSVQRTAGPAGFWES 358
QY 361 DDNDNMETASONGKRYKOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
Db 359 DDNDNMETASONGKRYKOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 418
QY 421 QAHVSSSNMAEFHASSHTWHTELTKTTDR 449
Db 419 QAHVSSSNMAEFHASSHTWHTELTKTTDR 447
```

RESULT 7  
P95564 PRELIMINARY; PRT; 447 AA.

```
AC P95564;
DT 01-MAY-1997 (TremBrel. 03, Created)
DT 01-MAY-1997 (TremBrel. 03, Last sequence update)
DT 01-JUN-2001 (TremBrel. 17, Last annotation update)
DE ISPalpA 2NT.
GN NTPAC.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J542;
RX MEDLINE=97128768; PubMed=8973308;
RA Parales J.V., Kumar A., Parales R.E., Gibson D.T.;
RT "Cloning and sequencing of the genes encoding 2-nitrotoluene
RT dioxygenase from Pseudomonas sp. J542.";
RL Gene 181:57-61(1996).
DR EMBL: U49504; AAB40383.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00335; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIPOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 447 AA; 49485 MW; ICBDE232E528E3BD CRC64;
```

Query Match 85.2%; Score 2051; DB 2; Length 447;  
Best local similarity 83.7%; Pred. No. 1,7e-149;  
Matches 376; Conservative 34; Mismatches 37; Indels 2; Gaps 1;  
QY 1 MNYNKILVSESGLSQKHLIHGDELFQHELTIFARNWMLFTHDSLIPADGVYVYAKMG 60  
Db 1 MSYQN--LVSEAGLTQKHLIHGDELFQHELTIFARNWMLFTHDSLIPSPGDYVYAKMG 58  
QY 61 IDEVIVSRQNDGSIIRAFILNVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELQSVPE 120  
Db 59 VDEVIVSRQNDGSIIRAFILNVCRRHGRKTLVHAEGNAKGFVCSYHGWGSGNGELQSVPE 118

[illegible]

RESULT	8			
Q8RTL4				
ID	Q8RTL4	PRELIMINARY;	PRT;	447 AA.
AC	Q8RTL4;			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Oxygenase-alpha NBD0.			
GN	NB2AC.			
OS	Comamonas sp. JS765.			
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas			
OX	NCBI_TaxID=58226;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JS765;			
RX	MEDLINE=21681044; PubMed=11823201;			
RA	Liesner D.J., Johnson G.R., Patales R.E., Spain J.C., Gibson D.T.;			
RT	"Molecular Characterization and Substrate Specificity of Nitrobenzene			
RL	Dioxygenase from Comamonas sp. Strain JS765.";			
RL	Appl. Environ. Microbiol. 68:634-641(2002).			
SQ	EMBL; AF379638; AAL76202.1; -			
SEQUENCE	447 AA; 49556 MW; 4CA6A6199A37DF3A CRC64;			

Query Match	84.0%;	Score 2022;	DB 2;	Length 447;
Best Local Similarity	82.0%;	Pred. No. 3e-147;		
Matches 368;	Conservative 40;	Mismatches 39;	Indels 2;	Gaps 1;

[illegible]

Qy	301	FPN\$M\$LTCSGVFKVWNPIDANTIEWTIAVEVDMPEDLKRRLABDVQNTAGPAGWES	360
		: : : : :	
Db	299	FPN\$S\$FLTGSAPRVNPEIDENTIEWTIAVEKDEMPEDLKRRLADAVQRSIGPAGWES	358
Qy	361	DDDNM\$ETASONGKRYOSRSDLL\$NIGFEGDYGDVAVYGVYKSAIGET\$TRGFPRAY	420
		: : : : :	
Db	359	DDNEM\$ETASONGKRYOS\$NIDQ\$IASIGFEGDYGDCEYGVYKSAIGET\$YRGFPRAY	418
Qy	421	QAHVSS\$NMAFEH\$AST\$WHT\$ELKTTDR	449
		: : : : :	
Db	419	QAH\$SS\$NMAFEH\$AS\$RWHT\$ELKTTDR	447

RESULT 9			
ID	Q45695	PRELIMINARY:	PRT; 451 AA.
AC	Q45695;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	ISP-alpha.		
GN	DNTAC.		
OS	Burkholderia sp. (strain RASC).		
OC	Bacteria; Proteobacteria; beta subdivision; Burkholderia group;		
OC	Burkholderia.		
OX	NCBI_TaxID=69003;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=DNT;		
RX	MEDLINE=96326338; PubMed=8759657;		
RA	Suen W.C., Hargler B.E., Spain J.C.;		
RT	"2,4-Dinitrotoluene dioxygenase from Burkholderia sp. strain DNT:		
RT	similarity to naphthalene dioxygenase.";		
RL	J. Bacteriol. 178:4926-4934 (1996).		
DR	EMBL; U62430; AAB09766.1; -.		
DR	HSSP; P23094; INDO.		
DR	InterPro; IPR001281; Rleske.		
DR	InterPro; IPR001663; Ring_hydroxyl_A.		
DR	Pfam; PF00355; Rleske; 1.		
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.		
DR	PRINTS; PR00090; RINGDIOXGNASE.		
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.		
SO	SEQUENCE 451 AA: 49827 MW: 726796C29GCF9A10 CRC64:		

Query Match	81.7%;	Score 1968.5;	DB 2;	Length 451;
Best Local Similarity	80.0%;	Pred. No. 3.9e-14;		
Matches 359;	Conservative 41;	Mismatches 46;	Indels 3;	Gaps 2

Oy	1	MNNKNTLSSSGISOKRLIHGDEEHOHEKTLFAANNMLFLTHDSLTPAPAGDVITAKMG	60
		6 MSYON--TVSBAGLITQKHLITGDLELFOHEKTLFAANNMLFLTHDSLTPSPGDVYKXMG	63
Oy	61	IDEVYVSRONDGSIRAFVNCVRHKGTLVSEACNAGKFCVSYHMGFCFSGNGELQSYFE	120
		64 VDEVYVSRONDGSIRAFVNCVRHKGTLVDAEAGNAGFCVCGHYMGYSGNGELQSYFFE	123
Oy	121	KDLYGSELNKKCLGIEKVARVESEFHGELYGCFDQEAEPILMDYLGDAAMYLEPMKHSGL	180
		124 KELYGDAITKKKCLGIEKVPRIESEHGFYLGCFDAEAPILMDYLGDAAMYLEPMKHSGL	183
Oy	181	ELVGEPPGKVYITKANMKAPAEVNFVDAYHVGWTHASLSRGESEJTSISLAGNAPLPEBGAGL	240
		184 ELVGEPPAVYVVKWGMKFAEENFVDIYHIGWTHASISLRAGAIAPLAGAAMPPEBTGL	243
Oy	241	QMTSKYSGMGVLMDGYSGVHSAOLVPELMAFGAKQDERLNKEIGDVRARITYSHLNCY	300
		244 QATTKYSGISVLDAYSQVSAOLVPELMAFGAKQEKELAKELGIDVRARITYRQVNVGT	303
Oy	301	FPNNMSLTCGVFVYVWNPIDANTAEVWTVYIVEGDMPEDLKRRLADSVORTAGPAGWES	360
		304 FPNCFCLTGAGVEVFNEDIDENTIEMVTLAVEGDMPEDLKRRLADDAQNSTGPAGWES	363
Oy	361	DDNNMFEASONGKRYKYSRDSLLSNLGFGEVDYGDVAIVPGVYKSAIGETSYRGFRAV	420

```
||||| ||| ||||| :|||||:||||| ||||| ||| :|:|||||
Db 364 DDNDNM-VLSQNAKKYOSNSDLADLGFGRDYGDECTPGVVSARSARETHRGFRAY 422
QY 421 QAHVSSNAEFAHSSWHTELKTDR 449
Db 423 QAHISSNAEFAHSSWHTELKTDR 451

RESULT 10
Q9ETK2 PRELIMINARY; PRT; 277 AA.
ID Q9ETK2
AC Q9ETK2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
RT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2 OR NAHAC.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-21DINH, PRIMN1, AND 31A2NH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306438; AAC25698.1; -.
DR EMBL; AF306432; AAC25692.1; -.
DR EMBL; AF306436; AAC25696.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDI_OXNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW dioxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30094 MW; BECFB5A379546DB CRC64;

Query Match 61.9%; Score 1490; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.7e-106;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVYSRONDSIRAFNLVNCRRHRTLVSVENAKKGFVCSYHGMFGSNGELQ 115
Db 1 TAKMGIDEIVYSRONDSIRAFNLVNCRRHRTLVSVENAKKGFVCSYHGMFGSNGELQ 60
QY 116 SVPEKDLVYGSNLKCKGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFK 175
Db 61 SVPEKDLVYGSNLKCKGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFK 120
QY 176 HSGGLELVGPPGKVIYKANKKAPAEFNGDAYHYGWTASSLRSGESIFSSLAGNAALPP 235
Db 121 HSGGLELVGPPGKVIYKANKKAPAEFNGDAYHYGWTASSLRSGESIFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGSGVYLMDSYGVHSADLVPELMAFGAGKOERLNKEIGDVRARIYRSH 295
Db 181 EGAGLQMTSKYSGSGVYLMDSYGVHSADLVPELMAFGAGKOERLNKEIGDVRARIYRSH 240
QY 296 LNCVTFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 332
Db 241 LNCVTFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 277

RESULT 11
Q9F5S4 PRELIMINARY; PRT; 277 AA.
ID Q9F5S4
AC Q9F5S4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
```

```
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-STIIASAL;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306440; AAC25700.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDI_OXNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW dioxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30094 MW; FADPDA5D37E241AB CRC64;

Query Match 61.8%; Score 1486; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 1.7e-106;
Matches 276; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVYSRONDSIRAFNLVNCRRHRTLVSVENAKKGFVCSYHGMFGSNGELQ 115
Db 1 TAKMGIDEIVYSRONDSIRAFNLVNCRRHRTLVSVENAKKGFVCSYHGMFGSNGELQ 60
QY 116 SVPEKDLVYGSNLKCKGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFK 175
Db 61 SVPEKDLVYGSNLKCKGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFK 120
QY 176 HSGGLELVGPPGKVIYKANKKAPAEFNGDAYHYGWTASSLRSGESIFSSLAGNAALPP 235
Db 121 HSGGLELVGPPGKVIYKANKKAPAEFNGDAYHYGWTASSLRSGESIFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGSGVYLMDSYGVHSADLVPELMAFGAGKOERLNKEIGDVRARIYRSH 295
Db 181 EGAGLQMTSKYSGSGVYLMDSYGVHSADLVPELMAFGAGKOERLNKEIGDVRARIYRSH 240
QY 296 LNCVTFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 332
Db 241 LNCVTFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 277

RESULT 12
Q9F5T3 PRELIMINARY; PRT; 277 AA.
ID Q9F5T3
AC Q9F5T3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
RT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 191IDNH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139738;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-191IDNH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
```

```

DR EMBL: AF306426; AAG25686.1; -.
DR HSP; P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PR00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Dioxxygenase.
KW NON_TER
FT
SQ
SEQUENCE 277 AA: 30064 MW: DB47868EB6D525A CRC64:

Query Match
Best Local Similarity 61.7%; Score 1486; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRONDGSTRAFNLVNCRRHGKTLVSEVAGNAKGFCVSHGWFSGNGELQ 115
DB 1 TAKMGIDEIVSRONDGSTRAFNLVNCRRHGKTLVSEVAGNAKGFCVSHGWFSGNGELQ 60

QY 116 SVPEKDLVGEISLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 175
DB 61 SVPEKDLVGEISLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 120

QY 176 HSGGLELVGPPGKVVYIKANMKAPAEVFGDAYHVGWTHASSLRGSEIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVYIKANMKAPAEVFGDAYHVGWTHASSLRGSEIFSSLAGNAALPP 180

QY 236 EGAGLOMTSKYSGMGVLMDSYGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSH 295
DB 181 EGAGLOMTSKYSGMGVLMDSYGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSH 240

QY 296 LNCITVFPPNNSMLTCSGVFKVWNPIDANTTEVNTAIV 332
DB 241 LNCITVFPPNNSMLTCSGVFKVWNPIDANTTEVNTAIV 277

RESULT 13
Q9F5S8 PRELIMINARY: PRT: 277 AA.
AC O9F5S8.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Naphthalene dioxxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. PR3MN2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR3MN2.
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF306434; AAG25694.1; -.
DR HSP; P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Dioxxygenase.
KW NON_TER
FT
SQ
SEQUENCE 277 AA: 30064 MW: 6BFF83D32F04CF03 CRC64:

Query Match
Best Local Similarity 61.6%; Score 1484; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRONDGSTRAFNLVNCRRHGKTLVSEVAGNAKGFCVSHGWFSGNGELQ 115
DB 1 TAKMGIDEIVSRONDGSTRAFNLVNCRRHGKTLVSEVAGNAKGFCVSHGWFSGNGELQ 60

QY 116 SVPEKDLVGEISLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 175
DB 61 SVPEKDLVGEISLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 120

QY 176 HSGGLELVGPPGKVVYIKANMKAPAEVFGDAYHVGWTHASSLRGSEIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVYIKANMKAPAEVFGDAYHVGWTHASSLRGSEIFSSLAGNAALPP 180

QY 236 EGAGLOMTSKYSGMGVLMDSYGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSH 295
DB 181 EGAGLOMTSKYSGMGVLMDSYGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSH 240

QY 296 LNCITVFPPNNSMLTCSGVFKVWNPIDANTTEVNTAIV 332
DB 241 LNCITVFPPNNSMLTCSGVFKVWNPIDANTTEVNTAIV 277

RESULT 14
Q9F5S9 PRELIMINARY: PRT: 277 AA.
AC O9F5S9.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Naphthalene dioxxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 8IDINH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8IDINH.
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF306433; AAG25693.1; -.
DR HSP; P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Dioxxygenase.
KW NON_TER
FT
SQ
SEQUENCE 277 AA: 29995 MW: AADA9B5A378FF6DB CRC64:

Query Match
Best Local Similarity 61.6%; Score 1483; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRONDGSTRAFNLVNCRRHGKTLVSEVAGNAKGFCVSHGWFSGNGELQ 115
DB 1 TAKMGIDEIVSRONDGSTRAFNLVNCRRHGKTLVSEVAGNAKGFCVSHGWFSGNGELQ 60

QY 116 SVPEKDLVGEISLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 175
DB 61 SVPEKDLVGEISLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 120

QY 176 HSGGLELVGPPGKVVYIKANMKAPAEVFGDAYHVGWTHASSLRGSEIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVYIKANMKAPAEVFGDAYHVGWTHASSLRGSEIFSSLAGNAALPP 180

QY 236 EGAGLOMTSKYSGMGVLMDSYGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSH 295
DB 181 EGAGLOMTSKYSGMGVLMDSYGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSH 240

```

Db 181 EGAGLQMTSKYSGMGVLMDSYGVSHSADLVPFLMARGAKQEGLNKEIGDVRARIYRSH 240  
QY 296 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 332  
Db 241 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 277

## RESULT 15

Q9F5S3 PRELIMINARY; PRT; 277 AA.  
AC Q9F5S3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 1 (Fragment).  
GN NAHAC1.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=511ANH;  
RA Ferrero M.A., Lalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in  
RT Pseudomonas strains from West Mediterranean Sea";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF306441; AAC25701.1; -.  
DR HSSP: P23094; INDO.  
DR InterPro: IPR001281; Rleske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rleske; 1.  
DR Pfam: PF00648; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIOXGNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
FT NON\_TER 1  
FT NON\_TER 277  
SQ SEQUENCE 277 AA; 30062 MW; 82C78D865A92D0AB CRC64;

Query Match 60.3%; Score 1451; DB 2; Length 277;  
Best Local Similarity 96.8%; Pred. No. 1.2e-103;

Matches 268; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRONCSIRAFILNVCRRKTLVSVAGNAKGFVCSYHGWFGSNGELQ 115  
Db 1 TAKMGIDEVIVSRQSDSIRAFILNVCRRKTLVNAEAGNAKGFVCSYHGWFGSNGELQ 60  
QY 116 SVPEKDLVGEISLNKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFK 175  
Db 61 SVPEKEIVGESLNKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFK 120  
QY 176 HSGGELVGPFGKVVYIKANKKAPENFVGDAYHYGWTTHASSLRSGESIFSSLAGNAALPP 235  
Db 121 HSGGELVGPFGKVVYIKANKKAPENFVGDAYHYGWTTHASSLRSGESIFSSLAGNAALPP 180  
QY 236 EGAGLQMTSKYSGMGVLMDSYGVSHSADLVPFLMARGAKQEGLNKEIGDVRARIYRSH 295  
Db 181 EGAGLQMTSKYSGMGVLMDSYGVSHSADLVPFLMARGAKQEGLNKEIGDVRARIYRSH 240  
QY 296 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 332  
Db 241 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWTYAIV 277

Search completed: January 27, 2003, 09:02:20  
Job time : 26.8333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:47:32 ; Search time 7.16667 Seconds  
(without alignments)  
2598.540 Million cell updates/sec

Title: US-09-843-250-33

Sequence: 2408  
1 MNYNKILVSESGLSQKHLI.....AEFHASSTWTELTKTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2402	99.8	449	1	NDOB_PSEPU
2	2316	96.2	449	1	NDOB_PSEFL
3	2217	92.1	449	1	NDOB_PSEAE
4	735	30.5	453	1	HCAE_ECOLI
5	650.5	27.0	450	1	BED1_PSEPU
6	648.5	26.9	450	1	TOD1_PSEPU
7	615.5	25.6	458	1	BPHA_BURCE
8	615	25.5	457	1	BPHA_COMTE
9	588.5	24.4	448	1	BNZA_PSEPS
10	585	24.3	458	1	BPA1_PSESI
11	410.5	17.0	454	1	XYRX_PSEPU
12	398	16.5	461	1	BEWA_ACICA
13	266	10.2	374	1	YEAW_ECOLI
14	239.5	9.9	442	1	CHMO_AMATR
15	233	9.7	438	1	CHMO_ATTRHO
16	224	9.3	446	1	CHMO_BETVU
17	223	9.3	439	1	CHMO_SPIOL
18	216	9.0	422	1	CHMO_ARATH
19	164	6.8	439	1	PHI3_PSEPU
20	118.5	4.9	354	1	VANA_PSEPS
21	110	4.6	329	1	VANA_PSEPS
22	110	4.6	432	1	CBAA_COMTE
23	109	4.5	426	1	PMP6_CHLPN
24	106	4.4	409	1	POBA_PSEPS
25	104.5	4.3	543	1	7DPI_DROME
26	104.5	4.3	746	1	7DPI_DROME
27	104.5	4.3	3255	1	POLG_LAME
28	102.5	4.2	1411	1	Y297_HUMAN
29	101.5	4.2	1926	1	LPH_RABIT
30	100.5	4.1	3255	1	POLG_LAMO
31	98.5	4.0	468	1	PPAI_PICPA
32	97.5	4.0	847	1	ORP8_HUMAN
33	97				Q9PZF1 homo sapien

34	96	4.0	1693	1	POIN_HEVMY	Q04610 hepatitis e
35	95	3.9	917	1	SYI_STRAU	P41972 staphylococ
36	94.5	3.9	331	1	LDHA_RHIDE	O96W58 rhidophila
37	94.5	3.9	350	1	UNRI_HUMAN	O9Y3F4 homo sapien
38	94.5	3.9	452	1	F26_YEAST	P32604 saccharomyc
39	94.5	3.9	518	1	ATPA_ENTHR	P26679 enterococcu
40	94	3.9	424	1	OAT_YEAST	P07991 saccharomyc
41	94	3.9	1693	1	POIN_HEVBU	P29324 hepatitis e
42	93	3.9	420	1	DBE3_PYRKO	O59650 pyrococcus
43	92.5	3.8	428	1	GBA1_CANAL	P28868 candida alb
44	92.5	3.8	525	1	NCAP_MEASA	P35972 measles vir
45	92	3.8	405	1	DCP2_PEA	P51851 pisum sativ

## ALIGNMENTS

RESULT 1  
ID NDOB\_PSEPU STANDARD: PRT: 449 AA.  
AC P23094: 052124: 033461: 007830:  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene  
1,2-dioxygenase ISP alpha).  
GN NDOB OR NAHAC OR DOXB OR NAH43 OR NDOC2.  
OS Pseudomonas putida, and  
OS Pseudomonas sp. (Strain C18).  
OG Plasmid pDrg1, Plasmid NAH7, and Plasmid NPL1.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OX NCBI\_TaxID=303, 306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=89211973; PubMed=3243438;  
RA Kurkela S., Lehtvaeslahti H., Palva E.T., Teeri T.H.;  
RT "Cloning, nucleotide sequence and characterization of genes encoding  
naphthalene dioxygenase of Pseudomonas putida strain NCIB9816.";  
RL Gene 73:355-362(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX Paries J.V., Paries R.E., Kumar A., Gibson D.T.;  
RA Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C18;  
RX MEDLINE=94042852; PubMed=8226631;  
RA Denome S.A., Stanley D.C., Olson E.S., Young K.D.;  
RT "Metabolism of dibenzothiophene and naphthalene in Pseudomonas  
strains: complete DNA sequence of an upper naphthalene catabolic  
pathway.";  
RL J. Bacteriol. 175:6890-6901(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=G7 / ATCC 17485; PLASMID=NAH7;  
RX MEDLINE=93252277; PubMed=8486285;  
RA Simon M.J., Ostlund T.D., Saunders R., Ensley B.D., Suggs S.,  
RA Harcourt A.A., Suen W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;  
RT "Sequences of genes encoding naphthalene dioxygenase in Pseudomonas  
putida strains G7 and NCIB 9816-4.";  
RL Gene 127:31-37(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=BS202; PLASMID=NPL1;  
RA Bezorodnikov S.G., Boronin A.M., Tiedje J.M.;  
RT "Nucleotide sequences of genes encoding an upper pathway of  
naphthalene metabolism of NPL1 plasmid from Pseudomonas putida strain  
BS202.";  
RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.  
RN [6]

RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=ATCC 17484;  
RA Hamann C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=98298434; PubMed=9634695;  
RA Kauppi B., Lee K., Carredano E., Parales R.E., Glibson D.T., Eklund H.,  
Ramawamy S.;  
RT "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene  
1,2-dioxygenase".  
RL Structure 6:571-586 (1998).  
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
NAPHTHALENE DIHYDRODIOL.  
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-  
dihydronaphthalene-1,2-diol + NAD(+).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES  
CONVERSION OF NAPHTHALENE TO SALICYLATE. AND CATABOLISM OF  
DIBENZOFIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO  
1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO  
OXIDATION OF THE AROMATIC RING.  
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME  
SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON  
SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED  
OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA). AND ISP IS  
COMPOSED OF AN HEXAMER OF THREE LARGE ALPHA SUBUNITS (NDOB) AND  
THREE SMALL BETA SUBUNITS (NDOC).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
DIOXYGENASE ALPHA SUBUNIT FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M23914; AAB47591.1; -;  
DR EMBL: U49496; AAA92141.1; -;  
DR EMBL: M60405; AAA16125.1; -;  
DR EMBL: M83949; AAA25902.1; -;  
DR EMBL: AF010471; AAB62707.1; -;  
DR EMBL: AF004284; AAB61373.1; -;  
DR PIR: JS0071; JS0071.  
DR PIR: B49343; B49343.  
DR PDB: INDO; 23-MAR-99.  
DR InterPro: IPR001281; Rieseke.  
DR InterPro: IPR001663; Ring-hydroxyl\_A.  
DR Pfam: PF00848; Ring-hydroxyl\_A; 1.  
DR Pfam: PF00355; Rieseke; 1.  
DR PRINTS: PR00090; RINGDIOXGNASE.  
DR PROSITE: PS00570; RING-HYDROXYL\_ALPHA; 1.  
DR Aromatic hydrocarbons catabolism: Oxidoreductase; Iron-sulfur; Iron;  
KM Dioxygenase; NAD; Plasmid; 3D-structure.  
FT METAL 81 81 IRON-SULFUR (2FE-2S).  
FT METAL 83 83 IRON-SULFUR (2FE-2S).  
FT METAL 101 101 IRON-SULFUR (2FE-2S).  
FT METAL 104 104 IRON-SULFUR (2FE-2S).  
FT METAL 208 208 IRON.  
FT METAL 213 213 IRON.  
FT METAL 362 362 IRON.  
FT VARIANT 4 4 N -> K (IN STRAIN G7).  
FT VARIANT 12 12 S -> F (IN STRAIN ATCC 17484).  
FT VARIANT 15 15 S -> T (IN STRAIN G7).  
FT VARIANT 32 32 K -> R (IN STRAIN G7).  
FT VARIANT 50 50 A -> S (IN STRAIN G7).  
FT VARIANT 70 70 N -> S (IN STRAIN G7).  
FT VARIANT 90 90 SV -> NA (IN STRAIN G7).  
-----

FT VARIANT 122 122 D -> E (IN STRAIN G7).  
FT VARIANT 173 173 M -> I (IN STRAIN G7).  
FT VARIANT 225 225 S -> A (IN STRAIN G7).  
FT VARIANT 225 225 S -> C (IN STRAIN BS202).  
FT VARIANT 232 232 A -> V (IN STRAIN G7).  
FT VARIANT 275 275 A -> S (IN STRAIN G7).  
FT VARIANT 391 391 E -> K (IN STRAIN G7).  
FT VARIANT 421 421 Q -> R (IN STRAIN ATCC 17484).  
FT VARIANT 434 434 H -> D (IN STRAIN G7).  
FT SEQUENCE 449 AA; 49607 MW; 1FD2F4229684F7A8 CRC64;  
Query Match 99.8%; Score 2402; DB 1; Length 449;  
Best Local Similarity 99.8%; Pred. No. 1.9e-181;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 MNYNKKILVSSGSLQKRLIHGDELFQHEKLTFFARNWMLFTRHSLIPAPGYVAKMG 60  
DB 1 MNYNKKILVSSGSLQKRLIHGDELFQHEKLTFFARNWMLFTRHSLIPAPGYVAKMG 60  
OY 61 IDEVIVSRONGSIRAFPLNVCRRHGKTLVSVEAGNAGFVCSYHGNGFSGNGLQSVPE 120  
DB 61 IDEVIVSRONGSIRAFPLNVCRRHGKTLVSVEAGNAGFVCSYHGNGFSGNGLQSVPE 120  
OY 121 KDLYGESLNKKCLGLEYARVESFHGFIYGCDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
DB 121 KDLYGESLNKKCLGLEYARVESFHGFIYGCDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
OY 181 ELVGPGRGVYVIRKAWKKAFAENFVGDAYHVGWTHASSIRSGSEFSSISLAGNAALPREGAGL 240  
DB 181 ELVGPGRGVYVIRKAWKKAFAENFVGDAYHVGWTHASSIRSGSEFSSISLAGNAALPREGAGL 240  
OY 241 QMTSKYSGMGVLMDGYSGVHSADLVEPLMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMDGYSGVHSADLVEPLMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300  
OY 301 FPNNSMLTCSGVFFVYVNPIDANTTEVWTYALVEKMPEDIKRRLADSVQRTAPAGFWES 360  
DB 301 FPNNSMLTCSGVFFVYVNPIDANTTEVWTYALVEKMPEDIKRRLADSVQRTAPAGFWES 360  
OY 361 DDNDMETSASONGKRYGSRDLSNLGFGEDYVGDAVYGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDMETSASONGKRYGSRDLSNLGFGEDYVGDAVYGVYKSAIGETSYRGFYRAY 420  
OY 421 QAHVSSNMWAEFEHASSTWHTELTKRTDR 449  
DB 421 QAHVSSNMWAEFEHASSTWHTELTKRTDR 449  
RESULT 2  
NDOB\_PSEFL STANDARD; PRT; 449 AA.  
ID NDOB\_PSEFL  
AC 007824;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene  
1,2-dioxygenase ISP alpha).  
GN NDOB OR NDOC.  
OS Pseudomonas fluorescens.  
OG Plasmid.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxId=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 17483;  
RA Hamann C.;  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
NAPHTHALENE DIHYDRODIOL.  
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-



```

CC dhidronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF004283; AAB1370.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske; 1.
CC Pfam: PF00848; Ring_hydroxyl_A; 1.
CC PRINTS: PR00090; RINGDIOXGNASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxxygenase; NAD; Plasmid.
CC METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 208 208 IRON (BY SIMILARITY).
CC METAL 213 213 IRON (BY SIMILARITY).
CC METAL 362 362 IRON (BY SIMILARITY).
CC SEQUENCE 449 AA; 49512 MW; 186A7924AD437A1C CRC64;

Query Match 96.2%; Score 2316; DB 1; Length 449;
Best Local Similarity 95.5%; Pred. No. 1,1e-174;
Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 MNYNKKILVSESGLSQKHLHSGDELFOHELKTIFARWMLFTTHDSLIPAPGDVYATKMG 60
DB 1 MNYNKKILVSESGLSQKHLHSGDELFOHELKTIFARWMLFTTHDSLIPSPGDVYATKMG 60
QY 61 IDEVIVSRONDSIRAFNVNCRHKGTLVSVAGNAKGFVCSYHGMGSGNCELOSYPE 120
DB 61 IDEVIVSRONDSIRAFNVNCRHKGTLVSVAGNAKGFVCSYHGMGSGNCELOSYPE 120
QY 121 KDLYGESLNNKCLGLKEVAVRSEFSGFTYGCDFQDEAPPLMDYLDGAAYLEPMFHSGL 180
DB 121 KDLYGESLNNKCLGLKEVAVRSEFSGFTYGCDFQDEAPPLMDYLDGAAYLEPMFHSGL 180
QY 121 KELYGESLNNKCLGLKEVAVRSEFSGFTYGCDFQDEAPPLMDYLDGAAYLEPMFHSGL 180
DB 121 KELYGESLNNKCLGLKEVAVRSEFSGFTYGCDFQDEAPPLMDYLDGAAYLEPMFHSGL 180
QY 181 ELVGGPGKVVYIKANNKAPENFVGDAYHVGTWTHASSLSNGESIFSSLAGNAALPREGAGL 240
DB 181 ELVGGPGKVVYIKANNKAPENFVGDAYHVGTWTHASSLSNGESIFSSLAGNAALPREGAGL 240
QY 241 QMTSYSGMGVLMDSYGVSADLVPELMAFGAKORLNEKIDVYARITRSLNCTV 300
DB 241 QMTSYSGMGVLMDSYGVSADLVPELMAFGAKORLNEKIDVYARITRSLNCTV 300
QY 301 PFNNSMLTCSGVFKVWNPIDANTTEVTYALVEKMDPEDLKRRLADSVQRTAGPAGFWS 360
DB 301 PFNNSMLTCSGVFKVWNPIDANTTEVTYALVEKMDPEDLKRRLADSVQRTAGPAGFWS 360
QY 361 DDNDMMETASONGKRYQSRDLSLNLGEGEDVDYDAVYPGVVGSAIGETSYRGFYAY 420
DB 361 DDNDMMETASONGKRYQSRDLSLNLGEGEDVDYDAVYPGVVGSAIGETSYRGFYAY 420

```

```

DB 361 DDNDMMETASONGKRYQSRDLSLNLGEGEDVDYDAVYPGVVGSAIGETSYRGFYAY 420
QY 421 QAHVSSSMNAEEFHAHSWHTLTKTDDR 449
DB 421 QAHVSSSMNAEEFHAHSWHTLTKTDDR 449

RESULT 3
NDOB_PSEAE STANDARD: PRT: 449 AA.
ID NDOB_PSEAE
AC Q51494;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene
DE 1,2-dioxygenase ISP alpha).
GN NDOB OR PAHA3.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAK1;
RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,
RA Kiyohara H.;
RT "The molecular analysis of an NAH7-type gene cluster, pah, located on
RT the chromosome of Pseudomonas aeruginosa PAK1."
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL.
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) -> (1R,2S)-1,2-
CC dhidronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D84146; BAA12240.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske; 1.
CC Pfam: PF00848; Ring_hydroxyl_A; 1.
CC PRINTS: PR00090; RINGDIOXGNASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxxygenase; NAD.
CC METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 208 208 IRON (BY SIMILARITY).

```

FT METAL 213 213 IRON (BY SIMILARITY).  
 FT METAL 362 362 IRON (BY SIMILARITY).  
 SQ SEQUENCE 449 AA; 49715 MW; 35A189136722A21C CRC64;  
 Query Match 92.1%; Score 2217; DB 1; Length 449;  
 Best Local Similarity 89.8%; Pred. No. 6,7e-167;  
 Matches 403; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

OY 1 MNVNNKLVSESLQKHLHGDDELFOHLEKTFARNWMLFLTHDSLIPAPGDVYAKMG 60  
 1 MNVNNKLVSESLQKHLHGDDELFOHLEKTFARNWMLFLTHDSLIPAPGDVYAKMG 60  
 DB 1 MNVNNKLVSESLQKHLHGDDELFOHLEKTFARNWMLFLTHDSLIPAPGDVYAKMG 60  
 OY 61 IDEVIVSRONDGSIKRAFLVNCRRHGGKTLVSEVGNAGKFCVSHGWCFSNGELQSVPE 120  
 61 IDEVIVSRONDGSIKRAFLVNCRRHGGKTLVSEVGNAGKFCVSHGWCFSNGELQSVPE 120  
 DB 61 IDEVIVSRONDGSIKRAFLVNCRRHGGKTLVSEVGNAGKFCVSHGWCFSNGELQSVPE 120  
 OY 121 KDLYGSLNKKCLGKLEVAVAVESFHGTYGCFDQEARPLMDYLDGDAWYLEPMKHSGL 180  
 121 KDLYGSLNKKCLGKLEVAVAVESFHGTYGCFDQEARPLMDYLDGDAWYLEPMKHSGL 180  
 DB 121 KDLYGSLNKKCLGKLEVAVAVESFHGTYGCFDQEARPLMDYLDGDAWYLEPMKHSGL 180  
 OY 181 ELVGPFGKVYIKANMKAPAEVFGDAVHVGWTHASSLRGSEIFSSLAGNAALPPEGAGL 240  
 181 ELVGPFGKVYIKANMKAPAEVFGDAVHVGWTHASSLRGSEIFSSLAGNAALPPEGAGL 240  
 DB 181 ELVGPFGKVYIKANMKAPAEVFGDAVHVGWTHASSLRGSEIFSSLAGNAALPPEGAGL 240  
 OY 241 QMTSKYSGMGVLMDSYGVSHADLVPELMAFGAKOERLNKEIGDVARITYRSHLNCV 300  
 241 QMTSKYSGMGVLMDSYGVSHADLVPELMAFGAKOERLNKEIGDVARITYRSHLNCV 300  
 DB 241 QMTSKYSGMGVLMDSYGVSHADLVPELMAFGAKOERLNKEIGDVARITYRSHLNCV 300  
 OY 301 FPNNSMLTSGVGVKVNPDANTTEVTVAVIVEKMDPEDIKRRRLADSVORTPAGAGWES 360  
 301 FPNNSMLTSGVGVKVNPDANTTEVTVAVIVEKMDPEDIKRRRLADSVORTPAGAGWES 360  
 DB 301 FPNNSMLTSGVGVKVNPDANTTEVTVAVIVEKMDPEDIKRRRLADSVORTPAGAGWES 360  
 OY 361 DDNDNMTASQNGKTKYQSRDLSNLGFGEDYVGDVAVPGVVKSAIGETSYNGFYRAY 420  
 361 DDNDNMTASQNGKTKYQSRDLSNLGFGEDYVGDVAVPGVVKSAIGETSYNGFYRAY 420  
 DB 361 DDNDNMTASQNGKTKYQSRDLSNLGFGEDYVGDVAVPGVVKSAIGETSYNGFYRAY 420  
 OY 421 QAHVSSNNAAEFHASSTWTELTKTDR 449  
 421 QAHVSSNNAAEFHASSTWTELTKTDR 449  
 DB 421 QAHVSSNNAAEFHASSTWTELTKTDR 449

RESULT 4  
 HCAE\_ECOLI STANDARD; PRT; 453 AA.  
 AC 047139; P77590; P78203;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 3-phenylpropionate dioxygenase alpha subunit (EC 1.14.1.-)  
 GN (Dioxygenin alpha subunit).  
 GN HCAE OR PHC1 OR HCAA OR HCAAI OR DIGA OR B2538 OR 23809 OR ECS3404.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 OC NCBI\_TaxID=562, 83334;  
 OX NCBI\_TaxID=562, 83334;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RL Turlin E., Gasser F., Biyille F.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE-97426617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN (3)  
 RP SEQUENCE FROM N.A.

RC STRAIN-K12;  
 RX MEDLINE-97349980; PubMed-9205837;  
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,  
 RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,  
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata S., Horinouchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli  
 RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and  
 RT analysis of its sequence features.";  
 RL DNA Res. 4:91-113(1997).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE-1074935; PubMed-11206511;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamistis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN (5)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / RIMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shida T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- FUNCTION: PART OF MULTICOMPONENT 3-PHENYLPROPIONATE DIOXYGENASE.  
 CC CONVERTS 3-PHENYLPROPIONIC ACID (PP) INTO CIS-3-(3-CARBOXYETHYL)-  
 CC 3,5-CYCLOHEXADIENE-1,2-DIOL (PP-DIHYDRODIOL).  
 CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
 CC -1- PATHWAY: 3-phenylpropionic acid catabolism.  
 CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE  
 CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (HCAE AND HCAI), A  
 CC FERREDOXIN (HCAC) AND A FERREDOXIN REDUCTASE (HCAD).  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 CC DIOXYGENASE ALPHA SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z37966; CAAB6018.1; -;  
 DR EMBL: AE000340; AAC75591.1; -;  
 DR EMBL: D90883; BAA16433.1; -;  
 DR EMBL: D90884; BAA16441.1; -;  
 DR EMBL: AE005484; AAG57651.1; -;  
 DR EMBL: AP002562; BAB36827.1; -;  
 DR HSSP: P23094; INDO.  
 DR EcoGene: EGI3456; hcaE.  
 DR InterPro: IPR001281; Rieske.  
 DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
 DR Pfam: PF00355; Rieske; 1.  
 DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
 DR PRINTS: PR00090; RINGDIOXGNASE.  
 DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
 KM Dioxygenase; NAD; Complete proteome.  
 FT METAL 85 85 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 87 87 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 105 105 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT	METAL	108	108	IRON-SUFUR (2FE-2S) (BY SIMILARITY).
FT	METAL	213	213	IRON (BY SIMILARITY).
FT	METAL	218	218	IRON (BY SIMILARITY).
FT	CONFLICT	20	20	V -> A (IN REF. 1).
FT	CONFLICT	394	453	GHRAANSKICLEMGGOEKRDDGIPGIIITNYIFSETAARGM
FT	FT			YORMDLISSESMOVLDTATAYOOEVMK -> ATAATAAN
FT	FT			CVMKGLVRSKSAATAPLALLTIFISFOKLPVBCNTAAGPIF
FT	FT			(IN REF. 1)
SO	SEQUENCE	453 AA;	51109 MM;	02355BF5f7643FD CRC64;
	Query Match	30.5%;	Score 735;	DB 1; Length 453;
	Best Local Similarity	36.4%;	Pred. No. 2.3e-50;	
	Matches 165;	Conservative 64;	Mismatches 174;	Indels 50; Gaps 8
QY	20	INGDELEPOHEKLTFFARFWLTLFTDLSLIPAGDVTYAKMGIDDEVYVSQNOGSTAFELN 79		
DB	24	ITTDPIIYLOELERLFFRGWMLFLAHESQIPKGDGDFNTYMGEDAVVYVQKDGSIKAFLN 83		
QY	80	VCRHKGKTLVSEAGNANGFVCSYHGWCFSNGELOSVPFEKDLGSEISNKKCLGLEVA 139		
DB	84	QCRHRAMRVSYADCGNTRAFITCPYHGWSYGINGELLIDVLEPRAYPDQGLCKSHWGLENYP 143		
QY	140	KRESHGPIYGCDFDEAPPLMDYLDDAANYLEPMR-KHSGGLELVGPPGKVVYIKANKAP 198		
DB	144	CYESYKGLIFGWMDSAGLRDYLTDIAWYIDGMMDRREGTEIYGVQKWVINCWKPEP 203		
QY	199	AENFYGDAVHVWMTTHASSL-----RSGESIFSLAGNALP----- 234		
DB	204	AEQFASDQYHALFHSIASHAOVLGADDGSDKRLGD-----GQIARPYWETAKDALQFG 256		
QY	235	--PEBAGLOMTSKSYSGMGVLMDSYGSVHSAIDLVELMAFGAKOERLKEIGDVARARY 292		
DB	257	QDGHSGCEFFTEKPPANVWV--DGAVSSYRETYAE-----AEQRLGEVARLRL 303		
QY	293	RSHLNCTEFPNNNSMLTGSVGFVKWMPIDANTTEVYTAIVKEDMPEDLKRRLADSVQRTA 352		
DB	304	AGHNN--IFPILSWLNGTATLRAWPRRGPDOYEVAAFCITTDKAASDEVAAAFENSATRAF 361		
QY	353	GPAGFESDDDDNNMETASQNGKRYOSRSDLSLNGFEDGVYGDVAVYPGVGWSKAIGETS 412		
DB	362	GPAGLEQDDSENNWCEIQKLKLGHRARNSKICLEMGLQGERKRDDGIPGIT-NYIFSEIA 420		
QY	413	YRGFYRAYOAHVSSSNMAFEHASTTWITELTK 445		
DB	421	ARGMTOFMAADLISSESMQEVLDKTAAYOOEVMK 453		
	RESULT 5			
	BED1_PSEPU	STANDARD:	PRT:	450 AA.
ID	BED1_PSEPU	STANDARD:	PRT:	450 AA.
AC	Q07944;			
DT	01-NOV-1995 (Ref. 32, Created)			
DT	01-NOV-1995 (Ref. 32, Last sequence update)			
DT	16-OCT-2001 (Ref. 40, Last annotation update)			
GN	Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3).			
OS	BEC1.			
OS	Pseudomonas putida.			
OC	Plasmid pMT112.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=303;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.			
RC	STRAIN-MU2;			
RX	MEDLINE=93345820; PubMed=8344526;			
RA	Tan H.-M., Tang H.-Y., Joannou C., Abdel-Wahab N.H., Mason J.R.;			
RT	"The Pseudomonas putida M2 plasmid-encoded genes for benzene			
RT	dioxygenase are unusual in codon usage and low in G+C content."			
RL	Gene 130.33-39(1993).			
CC	-I- CATALYTIC ACTIVITY: Benzene + NADH + O(2) = cis-1,2-			
CC	dihydrobenzene-1,2-diol + NAD(+).			
CC	-I- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.			
CC	-I- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.			

```

CC      -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC      TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BEDC1 AND BEDC2), A
CC      FERREDOXIN (BEDB) AND A FERREDOXIN REDUCTASE (BEDA).
CC      -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC      DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL, AF148496; AAL1758.1; -.
DR      HSSP, P23094; INDO.
DR      InterPro, IPR001281; R1eske.
DR      InterPro, IPR001663; R1ng_hydroxyl_A.
DR      Pfam, PF00355; R1eske; 1.
DR      Pfam, PF00848; R1ng_hydroxyl_A; 1.
DR      PRINTS, PR00090; RNCDIOXGNASE.
DR      PROSITE, PS00570; R1NG_HYDROXYL_ALPHA; 1.
DR      KW      Aromatic hydrocarbon catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW      Dioxygenase; NAD; Plasmid.
FT      METAL           96      96      IRRN-SULFUR (2FE-2S) (BY SIMILARITY) .
FT      METAL           98      98      IRRN-SULFUR (2FE-2S) (BY SIMILARITY) .
FT      METAL          116     116      IRRN-SULFUR (2FE-2S) (BY SIMILARITY) .
FT      METAL          119     119      IRRN-SULFUR (2FE-2S) (BY SIMILARITY) .
FT      METAL          222     222      IRRN (BY SIMILARITY) .
FT      METAL          228     228      IRRN (BY SIMILARITY) .
SQ      SEQUENCE      450 AA; 51108 MW; 1ACD5EA6C4CF72C8 CRC64;

```

Query Match	Similarity	27.0%:	Score 650.5:	DB 1:	Length 450:
Best Local Match	34.6%:	Pred. No. 1e+43:			
Matches 148:	Conservative	69:	Mismatches 174:	Indels	Gaps 37:
					11:
QY	20	IGHDEELFOHELKTIFARNWLFLETFDLSLIPAGDGVTAQKMGIDEIVYSRONDGSTRALFN 79			
Db	35	IYTDLDLYQLELERYAFASWLLGHETHIRKQGVFTTYMGEDPVYVVRQKDAIAYPLN 94			
QY	80	VCRHNGKTLVSYEAGNAKGFVCSYHGMGFGSNGELQSVPEFKDLGESLNKKCLGLKE-- 137			
Db	95	QCRHNGMRICRSDAAGNAKAFCTSYHGMAAYDTAGNLINPYEAESFA-----CLDKREMS 148			
QY	138	--VAVSEFHGFIYVCEQOEAPRLMDYDGDAAWYLEPWF-KHSGLELVGPPGKVIYIAN 194			
Db	149	PIKAVEVYIKGLIFENWDEMLIDDTIYGEKKFYMDHLDPTREATEVYIPGIQKKVIECN 208			
QY	195	WKAPAEVNVGDAYHNGWTF-HASSLRSGESIFSSLAGNALPPEGALQMTSKYSGMGVYL 253			
Db	209	WKFPAEQQCSMDYHAGTIAHLISGIIAGLPEDLELDLA--PP-----KFGQIYAS 257			
QY	254	WDGY-SGVHSHAD-----LVPELMAF--GGAKOERLNKEIGDYR--ARIYSHLNCVTF 301			
Db	258	WGHSGSGEYIDDPNMLAMGPKVYSLTEGPAEKAERLGSIERGKIMLEHW--TVF 315			
QY	302	PNNSLTOSGYFKVWNPFLDANTTEVMTAIYEEKMPELCKRLADSVQRTGAPAGFWESD 361			
Db	316	PTCSFLPQVNTIRTWHPGPGNEVEVWATVYDADAPDDIKEEFRQTLRTESAGGVFQD 375			
QY	362	DNDNNETASQNGKKYOSRSDSLSLNIGGGEVYGDVAVPGVVGSAIGETVRCGRYAYQ 421			
Db	376	DGEENVLEIQLHLIRGHAKASRPFNAEMSGQIVYDNDPIYGRISNNVYSEEAARGLIYHML 435			
QY	422	AHVSSSNW 429			
Db	436	KMMTSPDW 443			
RESULT 6					
TOD1_PSEPUP					
ID TOD1_PSEPUP					
NC P13450:		STANDARD:		PRT:	450 AA.

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
GN Toluene 2,3-dioxygenase alpha subunit (EC 1.14.12.-).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC STRAIN=Fl;
RX MEDLINE=89359301; PubMed=2670929;
RA Zylstra G.J., Gibson D.T.;
RT "Toluene degradation by Pseudomonas putida Fl. Nucleotide sequence of
RT the todC12BAD genes and their expression in Escherichia coli.";
RL J. Biol. Chem. 264:14940-14946(1989).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: Toluene degradation: first step.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (TODC1 AND TODC2), A
CC FERREDOXIN (TODB) AND A FERREDOXIN REDUCTASE (TODA).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: J04996; AAA26005.1; -.
DR PIR: A36516; A36516.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rleske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00335; Rleske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 222 IRON (BY SIMILARITY).
FT METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 50944 MW; 038C80F197E3485D CRC64;

Query Match 26.9%; Score 648.5; DB 1; Length 450;
Best Local Similarity 34.4%; Pred. No. 1.5e-43;
Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;

OY 20 IHGDEELFOHELTIFARNMLFTLHSLIPAPGDYVAKMGIDEVIVSRQNDGIRAFLN 79
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 35 IYDEDELYQLELBRVAFARSWLLGHEHQIRKPGDYITTYMGEDPVVVYVROQDASIAVELN 94
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 80 VCHNRKTLVSAGNAGKGVCSHGNGFSGNSGFLQSVPEKDLGSLNKKCLGKE-- 137
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 95 QCHNRMRICRADAGNAKATCTSYHGWAYDTAGNLVNVVPEAESFA-----CLNKKRWS 148
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 138 --VARVESFHGTYGCPDOAPRLMDYLGAAYLEPMF-KHSGGLELVGPGKVIKAN 194
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 149 PLKARVETTYGLIFANNDENAVLDLTYLGEAKFMHMLDRTGTEAIRPVQKWVLPEN 208
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 195 WKARAFNVGDAYHVGWT-HASSLRSG--ESIFSLAGNALPPEGAGLQMTSKYSGSGM 251
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 209 WKFAABQFCSDMYHAGTSHLSGLAGLPEDL-----EMADLAPPTYGKQYRASMGHGS 263
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 252 VLMDGISGVHSDLVPELMAF--GGAKEORLNKEIGDVR--ARIYRSHLNCTVPPNNSML 307
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

DB 264 GFIYGDPLMLAIMGPRVTSYWTGEPASEKAERLGSVERGSKLWEHM--TVFPFCSEFL 321
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 308 TCSGVKWNPNIDANTFEVWTFYALVEKDPEDLKRRLDLSVOTRAGPAGWESDDNDNE 367
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 322 PCITVTRVHMRGPRNEVEVNAFTYVDADAPDDIKEFRROTTLRTFSAGVFEDDDENNV 381
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 368 TASQNGKKYQSRDSDLNLCFGEVDYGDVAVPGVKSATIGETSYRGFYRAYQAHVSSS 427
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 382 EIQHILRCHKARSRPFNAEMSMQTDVNDPVPGRISNNYSEARGLYAHMLRMWTFSP 441
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 428 NW 429
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 442 DW 443

RESULT 7
BPHA_BURCE
ID BPHA_BURCE STANDARD; PRT; 458 AA.
AC P37333;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
DE dioxygenase).
GN BPHA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB400;
RX MEDLINE=92234948; PubMed=1569021;
RA Erickson B.D., Mondello F.J.;
RT "Nucleotide sequencing and transcriptional mapping of the genes
RT encoding biphenyl dioxygenase, a multicomponent
RT polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain
RT LB400.";
RL J. Bacteriol. 174:2903-2912(1992).
RN [2]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RC STRAIN=LB400;
RX MEDLINE=96011369; PubMed=7592331;
RA Haddock J.D., Gibson D.T.;
RT "Purification and characterization of the oxygenase component of
RT biphenyl 2,3-dioxygenase from Pseudomonas sp. strain LB400.";
RL J. Bacteriol. 177:5834-5839(1995).
RN [3]
RP ERRATUM.
RA Haddock J.D., Gibson D.T.;
RL J. Bacteriol. 178:2158-2158(1996).
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Biphenyl 1,2fe-2S cluster and 1 iron atom per subunit.
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase
CC (Bphc) must be present to obtain activity.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M86348; AAB63425.1; -.
DR PIR: B41858; B41858.

```

```

DR HSP; P23094; 1NDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIPOXNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxigenase; NAD.
FT INIT MET 0
FT METAL 99 99 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 122 122 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 232 232 IRON (BY SIMILARITY).
FT METAL 238 238 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51382 MW; AFSBFLD05BAFDF4 CRC64;

Query Match 25.6%; Score 615.5; DB 1; Length 458;
Best Local Similarity 32.9%; Pred. No. 5,9e-41;
Matches 146; Conservative 69; Mismatches 186; Indels 43; Gaps 12;

OY 11 ESGLSQKHLIHGDELFQHELTIFARNWFLTHDSLIAPGDVYATKMGIDEVYISRON 70
DB 30 EKGLLDPR-IYADQSLYLELELVEFRSWLLLGHESHVETGDFLATYGEDEVYVVRQK 88
OY 71 DGSIRAFNLVCHRGKTVSVENAGNAKGFVCSYHGWGFGSNGELOSVPPEKDLGSLNK 130
DB 89 DKSIVFLNQGCHRGRIQRSDAGNAKFTCSYHGWAYIACKLVAVPRKEKFCCKKRG 148
OY 131 KCLGLKEV-----ARVESFHGFIYGCDEQAPPLMDYIGDAWYLEPKFKHS-GGLELV 183
DB 149 DC-GEFKAEWGEFLQARVATYKGLVEFAMVDVQAPDLETTYLGDRPYMDVMDLDTPTAGTVAI 207
OY 184 GPRGVKIVIKAKAPENVGVDAHYHG-WTHASSLSRSGSITSSLAGNALRPES--AAL 240
DB 208 GGMQKWVIFCNKRFABEQCSDMYHAGTTHTLSGLIAG-----IPPEMDLSQA 255
OY 241 QMTSKYSGMGVLMQDYSVSHADYLPPELLAFGAK-----OERLNKEID 286
DB 256 QLPETK-GNGFRAMGSHGSGWYVDERGSLAVMGPRVTOYTWEGRAELAEQRLGHTGMP 314
OY 287 VVARIRKSLNCTVPPNNMILTCGVEKYWNPIDANTTWTYAYAEKMPEDLKRRLAD 346
DB 315 VR-RMWGQHMR--TIPPTCSFLPTFNIRIMHPRGPEIEVMFAFTLVADADAPAIKEEYRR 371
OY 347 SVQRTAGPRGPFMSDNDMMETASQNGKKYQSRDLSNLGFGEDVYDAYYPGVYKGS 406
DB 372 HUIRNSAGGVFEDDGEWWEIQKGLRGYKAKSOPLNQMGIGRSQTHPDPFGNVG-Y 430
OY 407 AIGETSYRGVFRAYQAHVSSSNMA 430
DB 431 VYAEBAARGMYHHMMKMSPEPSMA 454

RESULT 8
BPHA_COMTE STANDARD: PRT; 457 AA.
AC 046372:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxigenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
DE dioxigenase).
GN BPHA.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B-356;
RX MEDLINE=97045812; PubMed=8890734;
RA Sylvestre M., Stirois M., Hurtubise Y., Bergeron J., Ahmad D.,

```

```

RA Shareck F., Barriault D., Guillemette I., Juteau J.M.;
RT "Sequencing of Comamonas testosteroni strain B-356-
RT biphenyl/chlorobiphenyl dioxigenase genes: evolutionary relationships
RT among Gram-negative bacterial biphenyl dioxigenases.";
RL Gene 174:195-202(1996).
CC -i- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -i- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (by
CC similarity).
CC -i- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -i- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase
CC (Bphg) must be present to obtain activity (BY similarity).
CC -i- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U47637; AAC44526.1; -.
CC DR HSP; P23094; 1NDO.
CC DR InterPro; IPR001281; Rieske.
CC DR InterPro; IPR001663; Ring_hydroxyl_A.
CC Pfam; PF00355; Rieske; 1.
CC Pfam; PF00848; Ring_hydroxyl_A; 1.
CC PRINTS; PR00090; RINGDIPOXNASE.
CC DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxigenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 457 AA; 51691 MW; D13FC0635FACBF5 CRC64;

Query Match 25.5%; Score 615; DB 1; Length 457;
Best Local Similarity 33.1%; Pred. No. 6,4e-41;
Matches 144; Conservative 73; Mismatches 172; Indels 46; Gaps 13;

OY 20 IHGDELFQHELTIFARNWFLTHDSLIAPGDVYATKMGIDEVYISRONDSIRAFIN 79
DB 39 IYADQSLYLELEFRGSRMLMGHETHIPKIGDYLTWYMGEDPYIMVRQKQSIKVFILN 98
OY 80 VQRHNGKTVSYEAGNAKGFVCSYHGWGFGSNGELOSVPFEEDLIGESLNKCKLGLKEV- 138
DB 99 QCRHRRGRIVRSDGNAKAFCTYTHGMWADYAGNLVNVPEFEKACDKKEGDC-GEKRD 157
OY 139 -----ARVESFHGFIYGCDEQAPPLMDYIGDAWYLEPMF-KHSGGLELVGPPGVYIK 192
DB 158 WGPLDARVETTYGVLFPANMDPPAPDLKTYLSDAFPMYMDLDRTEAGTAIGTQKWYIP 217
OY 193 ANWKAPAEENVGVDAHYHG-WTHASSLSRSGSITSSLAGNALPPEGAGLQMT-----SKYQ 247
DB 218 CNMKFAAEQFCSDMYHAGTMSHLSGLAG-----LPPE-----MDLQIQLSKNG 262
OY 248 SGMGLVMQY-----SGVHSADLVPELMAF--GGAQOEALNKEIGDYR-ARIYRSL 296
DB 263 NQFRSAGHGAGWFTINDSILLVAVPRITQYWTQGPAAEAAARRVDPDLTIDHFGOHM 322
OY 297 NCTVPPNNMILTCGVEKYWNPIDANTTWTYAYAEKMPEDLKRRLADSVQRTAGPRAG 356
DB 323 --TVPTFCFLGINTIRKWHPRGNEVEYMAFVLVDADAPADIDKDEPRLQIRKIPNAG 380
OY 357 FWESDNDMMETASQNGKKYQSRDLSNLGFGEDVYDAYYPGVYKSA--IGETSYR 414
:| | | | : : : : : | : | | | |

```

```
Db      381 VPEODDGEWVWEIQRYMRGRHKAQKSTSLACKMGLNVPKNKNPVR---GCTATVYVEEAR 437
Oy      415 GFYRAYQAHVSSSNW 429
         | : : : : :
Db      438 GMYHHWSRMSEPSW 452

RESULT 9
BPHA_PSEPS
ID BPHA_PSEPS STANDARD: PRT: 458 AA.
AC Q52028;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (BC 1.14.12.18) (biphenyl 2,3-dioxygenase).
GN BPHA OR BPBH1.
OS Pseudomonas pseudoalcaligenes.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxId=330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KF707;
RX MEDLINE=92165849; PubMed=1537863;
RA Taira K., Hirose J., Hayashida S., Furukawa K.:
RT "Analysis of bph operon from the polychlorinated biphenyl-degrading strain of Pseudomonas pseudoalcaligenes KF707.";
RL J. Biol. Chem. 267:4844-4853(1992).
CC -I- CATALYTIC ACTIVITY: Biphenyl + NMDH + O(2) = (2R,3S)-3-phenylcyclohexa-3,5-diene-1,2-diol + NMD(+).
CC -I- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By similarity).
CC -I- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway; first step.
CC -I- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase (Bphg) must be present to obtain activity (By similarity).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M83673; AAA25743.1; -.
DR HSBP; P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RNCDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron; Dioxxygenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51437 MW; A6123AD09F97E482 CMC64;

Query Match 25.5%; Score 615; DB 1; Length 458;
Best Local Similarity 33.1%; Pred. No. 6,5e+41;
Matches 147; Conservative 70; Mismatches 183; Indels 44; Gaps 13;

Oy      11 ESGLSQKHLLIHDELDFQHELTIFRRMMFLTHSHLPAPEDDYTAAMGIDEIVASQN 70
         |||||:::||||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

```

Db 31 EKKLLDPR - IYADQSLXLELELEAVFCFRSMULLLGHSHNVPERGDELATYMGDDPVVMRQK 89
Oy 71 DGSIRFALNCRHNRGKTLTVSEARNAKFGVCSYHGMFGSGNGLQSPFEKDLGESLAK 130
Db 90 DKSITVELNOCRHGRMCRISCDGNAKAFCTCSYHGMADVADIGKLVNVPFEKEAFCKEKG 149
Oy 131 KCLGIKEV-----ARESPHGFIVCCPQDPA,PLMLDVLGAAMVLEPMFHS- GGLLEY 183
Db 150 DC-GRDKAEMGPILOARATYKGLGVFANMVPVADLETYLGDARPRMDVMLDRPAGTVAI 208
Oy 184 GPPGVVYIKAMKAPAFNFEVDAYHNG- WTHASSLSRSGESIFSSLAGAALPEEG--AGL 240
Db 209 GGMQKWVYIPCMWKRKAQAQFCSDMYHAGTMSHLSGLIAG-----MPPEMDLSHA 256
Oy 241 QMTSKYSGSGMGLMDGYSGVHSADLVPELMARFAGK-----OERLKEIGD 286
Db 257 QVPTK- GNOFRAGWGSGFVDEFGMLAVMGPKVTVQYTEGPAADLAEORLGHMT- P 314
Oy 287 VRARIYSHLNCYTFEPNNSMLTCSGFYKYNPIDANTTEVMTYALVEXDPEDLKRLAD 346
Db 315 VR-RMFGQHM--SYEPFCSPFLPAINTIRWHPRGPEIEIWAFTLVADADAELKEEYRR 371
Oy 347 SVQRTAGPAGFWESEDDDNMETASONGKYYGSRSDLSLNFGEEDYDGAIVPGCVYGS 406
Db 372 HNITFGAGCGFVEDDOSENMVELQKRLRQYKKAQSLPNAQGLRSQTHGHPDFGANG-Y 430
Oy 407 AIGETSYRGFYRAYOAHVSSSNMA 430
Db 431 VYAEAAARGMYHHMRMSEPSWA 454

RESULT 10
BNZA_PSEPU STANDARD; PRT; 448 AA.
AC P08084;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (BC 1.14.12.3) (p1 subunit),
GN BNZA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
OX NCBI_TaxId=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BE-81;
RX MEDLINE=88032840; PubMed=3667527;
RA "Trie S., Doi S., Yorifuji T., Takagi M., Yano K.:
RT "Nucleotide sequencing and characterization of the genes encoding
RT benzene oxidation enzymes of Pseudomonas putida.",
RL J. Bacteriol. 169:5174-5179(1987).
CC -1- CATALYTIC ACTIVITY: Benzene + NADH + O(2) = cis-1,2-
CC dihydrobenzene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2P-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BNZA AND BNZB), A
CC FERREDOXIN (BNZC) AND A FERREDOXIN REDUCTASE (BNZD).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17904; AAA25735.1; -
DR PIR: A29830; A29830.
DR HSSP: P23094; INDO.

```

```

DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 222 IRON (BY SIMILARITY).
FT METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 448 AA; 50208 MW; 6D1DBD5B2F040BE7 C6C64;

Query Match 24.4%; Score 588.5; DB 1; Length 448;
Best Local Similarity 32.9%; Pred. No. 7.6e-39;
Matches 142; Conservative 68; Mismatches 176; Indels 45; Gaps 11;

QY 20 IHGDELFQHLKTFEAFNNLFLTHDSLIIPAGDYVTAKMGIDEYIVSRONDGSTRATLN 79
DB 35 IYTDLDLQLELERYFAKSWLLGHETQIRKPGDITTYMGEDPVVVRQKDASTAVELN 94
QY 80 VCRHGRKTLVSEAGNAGKFCVSYHGNGFSGNGLQSVPEFKDLYGESLNKCKLGKEV- 137
DB 95 QCRHGRMTCRADGNAKAFCTSYHGWAYDTAGNLVNPVYEESFA-----CLNKKRMS 148
QY 138 --VARVESFHGFTYGCFOEAPPLMDYIGDAWYLEPMF-KHSGGLELVGPPGVYIKAN 194
DB 149 PLKAVEETFKGIFANWENAVDLDITYGEAKFYMDHMLDTEAGTEAIPGVQKWIVICN 208
QY 195 WKAPAEVVGDAVHGWT-HASSLSRG--ESIFSSLAGNALLPREGALQMTSKVSGMG 251
DB 209 WKFAEOPCSDMYHAGTTSLSGLIAGLPED-----EMADAPPTVYKQIRASGCHGS 263
QY 252 VLMDGYSVHSDADVPELMF--GGAKOERLNKEIGDVR--ARIYRSHLNCVTFPPNSML 307
DB 264 GFYVDPMLMALMGPKVTSYWTBGPASEKAAERLGSVERSKMLVHEM--TVPTCSFL 321
QY 308 TCSGVFKWNPIDANTTEVYTYAIEKMDPEDLKRRLADSVO-----RTAGPAGFW 358
DB 322 PGINVTIRLASRAERGEVMAFTVVDADAPDDIKEEFRRARRTESPVAUSSRTGRTG-- 379
QY 359 ESDDMNETFASONGKXKQSRDLSNLGFGEDYVGDVYPGVYKSAIETSVRGYR 418
DB 360 -----SSNSTSCFATSRSPNMEMSMDOYVNDPVPYGRISNNVYSEEAARGLYA 430
QY 419 AYOAHVSSSNW 429
DB 431 HMLRMWTSPOW 441

RESULT 11
BPA1_PSES1 STANDARD; PRT; 458 AA.
AC 052438:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
DE dioxygenase).
GN BPA1.
OS Pseudomonas sp. (strain KKS102).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=307;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94324970; PubMed=8048958;
RA Fukuda M., Yasukouchi Y., Kikuchi Y., Nagata Y., Kimbara K.,
RA Horiuchi H., Takagi M., Iano K.,
RT Identification of the bphn and bphb genes of Pseudomonas sp. strains
RT KKS102 involved in degradation of biphenyl and polychlorinated

```

```

RT biphenyls.";
RL Biochem. Biophys. Res. Commun. 202:850-856(1994).
CC -1- CAPALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 Iron atom per subunit (By
CC similarity).
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1- SUBUNIT: Heterohexamer consisting of 3 BPA1 subunits and 3
CC BPA2 subunits. A ferredoxin (BPA3) and a ferredoxin reductase
CC (BPA4) must be present to obtain activity (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D17319; BAA04137.1; -.
CC HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 103 103 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 121 121 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 124 124 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 234 234 IRON (BY SIMILARITY).
FT METAL 240 240 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51876 MW; 128B8C3B2A3CFA C6C64;

Query Match 24.3%; Score 585; DB 1; Length 458;
Best Local Similarity 33.6%; Pred. No. 1.5e-38;
Matches 147; Conservative 65; Mismatches 174; Indels 52; Gaps 14;

QY 20 IHGDELFQHLKTFEAFNNLFLTHDSLIIPAGDYVTAKMGIDEYIVSRONDGSTRATLN 79
DB 40 IYADLDLQLELERYFAKSWLLGHETQIRKPGDITTYMGEDPVVVRQKDASTAVELN 99
QY 80 VCRHGRKTLVSEAGNAGKFCVSYHGNGFSGNGLQSVPEFKDLYGESLNKCKLGKEV- 138
DB 100 QCRHGRMTCRADGNAKAFCTSYHGWAYDTAGNLVNPVYEKAPCDKKEGDC-GFDKAD 158
QY 139 -----ARVESFHGFTYGCFOEAPPLMDYIGDAWYLEPMF-KHSGGLELVGPPGVYIK 192
DB 159 WGPLQARVETFKGIFANWDAEAPDLKTYLSDAPYMDVMDRTFAGTGVYGMQKWIVP 218
QY 193 ANWKAPAEVVGDAVHG--WTHASSLSRGESIFSSLAGNALLPREGALQMT-----SKYG 247
DB 219 CNWFFAAQCFSDWYHAGTMAHLSGVLS-----SLPPE--MDLTQYQMSKNG 263
QY 248 SGKGVLMGYSVHSDADVPELMFAGK-----QERLNKEIGDVVAR--IYRSHL 296
DB 264 SIFRAWGHGSGWFINDAALILMAVMGKITQYWTGPALEAKAKRLNOMTYQMFQGHM 323
QY 297 NCTVFPNNSMLTCSGVFKWNPIDANTTEV--WTVYAIKMDPEDLKRRLADSVORTAG 353
DB 324 --TVPTCSFLPGINTIRSWHPRGPNEVECCPSWSMPMR--PEDIKEEPRRONIRTFN 378
QY 354 PAGWESDDNNMETASONGKXKQSRDLSNLGFGEDYVGDVYPGVYKSA--IGET 411
DB 379 AGGTFEODDGENWELQGLGHRKAKSAPLCAQMGVLNPNKSNDFP--GKTAYVYAE 435
QY 412 SYRGFYRAYQAHVSSSNW 429

```





```

KW Dioxigenase; NAD.
FT METAL 95 95 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 97 97 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 118 118 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 224 224 IRON (BY SIMILARITY).
FT METAL 229 229 IRON (BY SIMILARITY).
SQ SEQUENCE 461 AA; 52228 MW; 52228 MW; CFC3247A3C4379 CRC64;

Query Match 16.5%; Score 398; DB 1; Length 461;
Best Local Similarity 27.5%; Pred. No. 7,5e-24;
Matches 130; Conservative 78; Mismatches 191; Indels 74; Gaps 21;

QY 5 NKLTV--SSGSLQKH-LIHGDELFQHEHLEKTFIFANMLFLHDSILIPAGDYVTKMGCI 61
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 16 DELVYNTETGEFKLHRSFTDQALFDLEMKITFECSNMWYLHESQIPNNNDYTTIYGR 75
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 62 DEVYSRQNDGSIKAFNLVNCRRHGRKTLVSVEAGNAKGFVCSYHGMGFGSNGELQSVPEK 121
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 76 OPTLIARNNGELINAMINACSHRGAQLCRHRRGNKTTYTCPEHGMFTNNSGKLKVKDPS 135
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 122 DL-YGSLMKK-CLGKEXARVSEFHGFTYCGFDEAPRLMYLGDAAVYLERPMKHS- 178
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 136 DAGYSDCFMDDSGHDKVAFRESEYKGFLEGLSDPVSPLQEFLETKIIMIVGQSDQ 195
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 179 GLELVGPKGVVYIKANWKAPEFVGDAYHVG--WTHASSL-----RSGESIFSSLG 229
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 196 GLEVLGAVSTYTYEGNMKLTLEN-GADGYHVASAHNNVYATYOHREKQAGPTIRMSAG 254
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 230 NALPREGAGLQMTSKYG--SGMGVLMDSYSGVSHADIVPELM---AEGAKOERLNKE 283
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 255 SWG--KHGG-----SYGFEHGHMLMTQGNEDPRNPKALEYTERKGAMSKWM--- 304
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 284 IGVRRARIRSHLNCVFEFNNSMITCSG-VFKVWNPIDANTIEVWTYALVE-KDMPEDLK 341
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 305 -----IERSR-NCLLYNVYILMDQFSGQIRLRLISVKNKEVYTYCIAPVGEAREARA 356
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 342 RRL-----ADSVOR---TAGPAGFESDNDNMNETASQNGKYYQSRSDL 383
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 357 RRIROYEDFFNMSGMATPDDLEFRACQAGYAGI-ELEND-----MCRSKMIWYCPDPA 411
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 384 LSNLGEDEVDYDAVYGVVAKSAIGETSYRGFRRAYQAHVSSMMAEFEHAS 436
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 412 ANEIG-----LKPASIGIKTEDEGLYLAQHGYLWLSMKQALAAKEFPAS 455
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14
YEAH_ECOLI STANDARD; PRT; 374 AA.
AC P76253;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative dioxigenase alpha subunit yeast (EC 1.14.1.-).
GN YEAM OR B1802 OR Z2845 OR ECS2511.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxId=562, 83334;
RN NCBI_TaxId=562, 83334;
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RT Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;

```

```

RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,
RA Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map."
RT DNA Res. 3:379-392(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RT DNA Res. 8:11-22(2001).
CC -|- CORFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -|- SUBUNIT: PROBABLY HETERODIMER OF YEAM AND YEAX.
CC -|- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AE000274; AAC74872.1; -.
DR EMBL, D90823; BAA15597.1; ALT_INIT.
DR EMBL, D90824; BAA15606.1; ALT_INIT.
DR EMBL, AE005403; AAG56791.1; -.
DR EMBL, AP002558; BAB35934.1; -.
DR Ecocore; EG13509; yeaw.
DR InterPro; IPR001281; Rleske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rleske; 1.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; FALSE NEG.
KM Hypothetical protein; Oxidoreductase; Iron-sulfur; Iron; Dioxigenase;
KM NAD; Complete proteome.
FT METAL 89 89 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 91 91 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 109 109 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 211 211 IRON (BY SIMILARITY).
FT METAL 216 216 IRON (BY SIMILARITY).
SQ SEQUENCE 374 AA; 42561 MW; BB5386ACA9585606 CRC64;

Query Match 10.2%; Score 246; DB 1; Length 374;
Best Local Similarity 26.6%; Pred. No. 5.1e-12;
Matches 98; Conservative 53; Mismatches 127; Indels 90; Gaps 19;

```

```

QY 23 DEELFOHEIKTIFARNWLEFLHDSILIPAGDYVTKMGIDEVYISRQNDGSIKAFNLVNCR 82
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 31 DONAFEHENFVFAKSWICVASHSELANANDVYTRREITIGESTIVLRCGRDKYLRAYVNC 90
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:55:37 : Search time 13.1667 Seconds  
(without alignments)  
3278.305 Million cell updates/sec

Title: US-09-843-250-33

Perfect score: 2408

Sequence: 1 MNYNKKILVSESGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: pir1:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2402	99.8	449	2 JS0071	naphthalene dioxyg
2	2402	99.8	462	2 S27632	naphthalene 1,2-di
3	2344	97.3	449	2 JN0644	naphthalene 1,2-di
4	2290	95.1	449	2 C55217	polycyclic aromati
5	2051	85.2	447	2 JC5352	2-nitrotoluene dio
6	955	39.7	459	2 T31134	naphthalene dioxyg
7	772	32.1	450	2 T31256	terminal oxygenase
8	735	30.5	453	2 A65031	biphenyl dioxygena
9	735	30.5	453	2 D91054	biphenyl dioxygena
10	735	30.5	453	2 G85898	biphenyl dioxygena
11	723.5	30.0	455	2 T31258	aromatic oxygenase
12	650.5	27.0	450	1 JN0812	benzene 1,2-dioxyg
13	648.5	26.9	450	1 A36516	toluene dioxygenas
14	637	26.5	461	2 S51757	biphenyl dioxygena
15	615.5	25.6	459	1 B41858	biphenyl dioxygena
16	615	25.5	457	1 JC4993	biphenyl dioxygena
17	615	25.5	458	1 A42409	biphenyl dioxygena
18	603	25.0	431	2 JN0908	carbazole dioxygen
19	588.5	24.4	448	1 A29830	benzene 1,2-dioxyg
20	585	24.3	458	2 JC2467	biphenyl dioxygena
21	450	18.7	469	2 T50943	dioxygenase D1A1,
22	417	17.3	471	2 T31281	benzoate 1,2-dioxy
23	410.5	17.0	454	2 A41659	benzoate 1,2-dioxy
24	407.5	16.9	455	2 E83332	toluene 1,2-dioxyg
25	393	16.3	461	2 S23477	probable benzoate
26	387.5	16.1	464	2 G83331	anthranilate dioxy
27	336.5	14.0	424	2 E83384	probable ring-hydr
28	330	13.7	426	2 T31278	biphenyl dioxygena
29	312.5	13.0	391	2 T31251	aromatic oxygenase

30	311.5	12.9	468	2 G97447	hypothetical prote
31	311.5	12.9	468	2 AH2665	ring hydroxylating
32	291	12.1	429	2 F82970	probable ring hydr
33	287	11.9	412	2 H95311	probable aromatic-
34	283	11.8	415	2 AG3320	benzoate 1,2-dioxy
35	269.5	11.2	404	2 H87635	Rieske 2Fe-2S fami
36	263	10.9	420	2 T31285	biphenyl dioxygena
37	249.5	10.4	374	2 AF0304	probable dioxygena
38	246	10.2	374	2 C85791	probable choline m
39	246	10.2	374	2 G90942	probable choline m
40	246	10.2	374	2 B64941	probable choline m
41	233.5	9.7	382	2 G70946	probable dioxygena
42	224	9.3	446	2 T14542	choline monooxygen
43	223	9.3	439	2 T09214	choline monooxygen
44	214.5	8.9	186	2 J05354	2-nitrotoluene dio
45	186	7.7	426	2 T08350	choline monooxygen

#### ALIGNMENTS

RESULT 1	JS0071	naphthalene dioxygenase (EC 1.14.12.-) ndob protein - Pseudomonas putida
C:Species:	Pseudomonas putida	
C:Date:	31-Mar-1992	#sequence_revision 31-Mar-1992 #text_change 20-Aug-1999
C:Accession:	JS0071	
R:Kurkela, S.; Leivaeslahti, H.; Palva, E.T.; Teeri, T.H.		
Gene 73, 355-362, 1988		
A:Title:	Cloning, nucleotide sequence and characterization of genes encoding naphthal	
A:Reference number:	JS0070; MUID:89211973; PMID:3243438	
A:Accession:	JS0071	
A:Molecule type:	DNA	
A:Residues:	1-449 <KUR>	
A:Cross-references:	GB:M23914; NID:G151392; PIDN:AB47591.1; PID:G151394	
C:Keywords:	2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein	
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>		
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status		
Query Match	99.8%	Score 2402; DB 2; Length 449;
Best Local Similarity	99.8%	Pred. No. 1.3e-183;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
OY 1	MNYNKKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLTPAPDVTAKMG 60	
DB 1	MNYNKKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLTPAPDVTAKMG 60	
OY 61	IDEVIVSRONDGSIARFLNVCNHRKTLVSVAGNAKGVCSYHGMGRSGNCELOSVPFE 120	
DB 61	IDEVIVSRONDGSIARFLNVCNHRKTLVSVAGNAKGVCSYHGMGRSGNCELOSVPFE 120	
OY 121	KDLVESLNKKKCLGKEVARVESFHGFIYGCDFDQAPPLMDLGAAYLLEPMFHSGGL 180	
DB 121	KDLVESLNKKKCLGKEVARVESFHGFIYGCDFDQAPPLMDLGAAYLLEPMFHSGGL 180	
OY 181	ELVGPGRKVVIRKANKKAPAEENVGDAYHVGTWTHASLSRSGESIFSSLAGNMALEPEGAGL 240	
DB 181	ELVGPGRKVVIRKANKKAPAEENVGDAYHVGTWTHASLSRSGESIFSSLAGNMALEPEGAGL 240	
OY 241	QMTSYSGSGMGVLMGYSGVHSADLVPELMAFGAKKQERLNKEIDVYARIRSHLNCIV 300	
DB 241	QMTSYSGSGMGVLMGYSGVHSADLVPELMAFGAKKQERLNKEIDVYARIRSHLNCIV 300	
OY 301	FPNNSMLTCSGVEFKVWNPIDANTTEVMTYAIYEKMPEDLKRRLADSVORTAGPAGFES 360	
DB 301	FPNNSMLTCSGVEFKVWNPIDANTTEVMTYAIYEKMPEDLKRRLADSVORTAGPAGFES 360	
OY 361	DDNDMETASONGKKYQSRDSDLNLGFGEDVDYDAVYGVGKSAIGETSYRGFYRAY 420	
DB 361	DDNDMETASONGKKYQSRDSDLNLGFGEDVDYDAVYGVGKSAIGETSYRGFYRAY 420	

Db 361 DDNDNMETASONGKKYOSRSDLSNLGFGEEDYVGDVYPGVWKSALGETSYRGFYRAY 420

Oy 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449

Db 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449

## RESULT 2

naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large chain  
S27632  
C:Species: Pseudomonas sp.  
C:Date: 06-Jan-1995 #sequence\_rev1sion 06-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S27632; B49343  
R:Denome, S.A.; Young, K.D.  
submitted to the EMBL Data Library, February 1992  
A:Description: Cloning and molecular characterization of genes involved in metabolism of  
A:Reference number: S27631  
A:Accession: S27632  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <DEN>  
A:Cross-references: EMBL:M60405  
J:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
J: Bacteriol. 175, 6890-6901, 1993  
A:Title: Metabolism of dibenzothioophene and naphthalene in Pseudomonas strains: complete  
A:Reference number: A49343; MUID:94042852; PMID:8226631  
A:Accession: B49343  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 14-462 <DE2>  
A:Cross-references: GB:M60405; NID:g151195; PIDN:AAA16125.1; PID:g294351  
A:Experimental source: strain C18  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:84-132/Domain: Rieske [2Fe-2S] homology <RSK>  
F:94,96,114,117/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 99.8%; Score 2402; DB 2; Length 462;

Best Local Similarity 99.8%; Pred. No. 1,4e-183;

Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MNNYNNKILVSESGISOKHLIHGDEBELFQHELKTIFFARNWMLFTHDSILIPAPGDVYTKMG 60  
Db 14 MNNYNNKILVSESGISOKHLIHGDEBELFQHELKTIFFARNWMLFTHDSILIPAPGDVYTKMG 73

Oy 61 IDEVIVSRONDGSIIRAFILVNCRRHKGTLVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
Db 74 IDEVIVSRONDGSIIRAFILVNCRRHKGTLVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 133

Oy 121 KDLYGSLNKKCLGKLEVARVESFHGFTYGCFOEAPPLMDYLGDAWYLEPMPKHSGL 180  
Db 134 KDLYGSLNKKCLGKLEVARVESFHGFTYGCFOEAPPLMDYLGDAWYLEPMPKHSGL 193

Oy 181 ELVGPFGKVIYIKANMKAPAEFNGDAVHGMTHASSLRSGESIFSSLAGNAALPPEGAGL 240  
Db 194 ELVGPFGKVIYIKANMKAPAEFNGDAVHGMTHASSLRSGESIFSSLAGNAALPPEGAGL 253

Oy 241 QMTSKYSGMGVLMDYSGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSHLNTCV 300  
Db 254 QMTSKYSGMGVLMDYSGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSHLNTCV 313

Oy 301 FPNNSMLTSCGVFKVWNPIDANTTEVWTYAIVEKDMPEDLKRLADSVORTFPGAGFWES 360  
Db 314 FPNNSMLTSCGVFKVWNPIDANTTEVWTYAIVEKDMPEDLKRLADSVORTFPGAGFWES 373

Oy 361 DDNDNMETASONGKKYOSRSDLSNLGFGEEDYVGDVYPGVWKSALGETSYRGFYRAY 420  
Db 374 DDNDNMETASONGKKYOSRSDLSNLGFGEEDYVGDVYPGVWKSALGETSYRGFYRAY 433

Oy 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449  
Db 434 QAHVSSNMAEFHASSTWHTELTKTTDR 462

## RESULT 3

JN0644  
naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large cha  
N:Alternate names: nahAC protein  
C:Species: Pseudomonas putida  
C:Date: 31-Dec-1993 #sequence\_rev1sion 31-Dec-1993 #text\_change 20-Aug-1999  
C:Accession: JN0644  
R:Simon, M.J.; Oslund, T.D.; Saunders, R.; Ensley, B.D.; Suggs, S.; Harcourt, A.; Su  
Gene 127, 31-37, 1993  
A:Title: Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida st  
A:Reference number: JN0640; MUID:93252277; PMID:8486285  
A:Accession: JN0644  
A:Molecule type: DNA  
A:Residues: 1-449 <SIM>  
A:Cross-references: GB:M83949; NID:g151384; PIDN:AA25902.1; PID:g151387  
C:Comment: This protein is a member of naphthalene dioxygenase multicomponent enzyme  
C:Genetics:  
A:Gene: nahAC  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2  
C:Keywords: 2Fe-2S; heterotetramer; metalloprotein; oxidoreductase; Rieske iron-sulfu  
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 97.3%; Score 2344; DB 2; Length 449;

Best Local Similarity 96.7%; Pred. No. 5.6e-179;

Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

Oy 1 MNNYNNKILVSESGISOKHLIHGDEBELFQHELKTIFFARNWMLFTHDSILIPAPGDVYTKMG 60  
Db 1 MNNYNNKILVSESGISOKHLIHGDEBELFQHELKTIFFARNWMLFTHDSILIPAPGDVYTKMG 60

Oy 61 IDEVIVSRONDGSIIRAFILVNCRRHKGTLVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
Db 61 IDEVIVSRONDGSIIRAFILVNCRRHKGTLVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120

Oy 121 KDLYGSLNKKCLGKLEVARVESFHGFTYGCFOEAPPLMDYLGDAWYLEPMPKHSGL 180  
Db 121 KELYGSLNKKCLGKLEVARVESFHGFTYGCFOEAPPLMDYLGDAWYLEPMPKHSGL 180

Oy 181 ELVGPFGKVIYIKANMKAPAEFNGDAVHGMTHASSLRSGESIFSSLAGNAALPPEGAGL 240  
Db 181 ELVGPFGKVIYIKANMKAPAEFNGDAVHGMTHASSLRSGESIFSSLAGNAALPPEGAGL 240

Oy 241 QMTSKYSGMGVLMDYSGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSHLNTCV 300  
Db 241 QMTSKYSGMGVLMDYSGVHSADLVPELMAFGAKOERLNKEIGDVRARIYRSHLNTCV 300

Oy 301 FPNNSMLTSCGVFKVWNPIDANTTEVWTYAIVEKDMPEDLKRLADSVORTFPGAGFWES 360  
Db 301 FPNNSMLTSCGVFKVWNPIDANTTEVWTYAIVEKDMPEDLKRLADSVORTFPGAGFWES 360

Oy 361 DDNDNMETASONGKKYOSRSDLSNLGFGEEDYVGDVYPGVWKSALGETSYRGFYRAY 420  
Db 361 DDNDNMETASONGKKYOSRSDLSNLGFGEEDYVGDVYPGVWKSALGETSYRGFYRAY 420

Oy 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449  
Db 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449

RESULT 4  
C55217  
polycyclic aromatic hydrocarbon dioxygenase (EC 1.14.12.-) iron-sulfur protein large  
C:Species: Pseudomonas putida  
C:Date: 05-May-1995 #sequence\_rev1sion 05-May-1995 #text\_change 20-Jun-2000  
C:Accession: C55217  
R:Takizawa, N.; Kaida, N.; Toriige, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara  
J. Bacteriol. 176, 2444-2449, 1994  
A:Title: Identification and characterization of genes encoding polycyclic aromatic hy  
A:Reference number: A55217; MUID:94209249; PMID:8157615  
A:Accession: C55217  
A:Status: preliminary



Db 186 LKSTLACNMKVPTEPNFVGYHVGWTHAALOMIGELAGLSGNRADMPFDDLGLOFTMR 245  
QY 246 YGSGMGVLMDOYSGVH--SADLVPELMAFGAKOERLKEIGDVARLYRSHLNTVPPN 303  
Db 246 HGHGFGILDNAATAIHVKROCYVYKLEETRGCIKKRPE----REKLYGHMWTSLPN 301  
QY 304 NSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMEDLKRRLADSVORTAGPAGESDN 363  
Db 302 CSFLYGTNFKRMHPRGPHIEVWYTMVPKNADTEFKRSIQREAIRSFGAGTLESDDG 361  
QY 364 DNNETAQNKKYOSRSDLSNLGFGEDVYG-DAVYPGVYKSAIGCTSTRGTFRAYQA 422  
Db 362 ENNSATAYNNNGNITTRGCRM--NSSMKDREGPPHYPGIVGSFICETSYRGYRFWQE 419  
QY 423 HVSSSNMAEFEEHASTW 439  
Db 420 MLDAPDMAIRANDDTW 436

## RESULT 7

T31256  
terminal oxygenase component large chain homolog - Sphingomonas aromaticivorans plasmid  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31256  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A:Accession: T31256  
A:Residues: 1-450 <ROM>  
A:Molecule type: DNA  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378397; PIDN:AA003980.1  
C:Genetics:  
A:Genome: plasmid pML1  
A:Note: Dphala  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 32.1%; Score 772; DB 2; Length 450;  
Best Local Similarity 36.1%; Pred. No. 1.2e-53;

Matches 157; Conservative 88; Mismatches 184; Indels 6; Gaps 6;

QY 1 MNVNNKILVESGLSQKHLHGDLEFQHELKTIFARNMLFLTHDSLIPAPGDVYTKMG 60  
Db 1 MNDSTADLVDSRTGRSRSITYASEDITRGELERIFGKWLFLVHTSQIPKGFDFRTFMG 60  
QY 61 IDEVIVSRQDGSIRAFNLVCRHNGKTLVSVEANAGFCVSYHGWFSGNGELQSYFFE 120  
Db 61 EDDVYIYRQDGSIRAFNLVCRHNGKTLVSVEANAGFCVSYHGWFSGNGELQSYFFE 120  
QY 121 KDLVGSLSNKKCLGLKEVAVESFHGFIYGCEDQAPRLMDYLGDAAWYLEP-MFKHSGG 179  
Db 121 NEAVYFELBTRFKGLLPVTYVAEYKGLVFGCMWANSPLDYLDAKFEFLDWMADMPG 180  
QY 180 LELVGPGRKVVIAKMKAPENFVGDAVYHGMTHASSLRSGEST-FSSSLA-GNALPREG 237  
Db 181 SALVGETQKAVLDTNMVLPVENVCGDGYHLGMAHAGMAAQAQSDMLTGLSVNGSGVDLDG 240  
QY 238 AGLOMTSKYSGMGVLMDOYSGVH--SADLVPELMAFGAKOERLKEIGDVAR-ITYRSHL 296  
Db 241 -GLSVAGMGHMYLSALDGVSGYAFYDPKPILEYLEANQTVLDRGLGEVNGROYWQAQV 299  
QY 297 NCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMEDLKRRLADSVORTAGPAG 356  
Db 300 NITIFPNLQDLPLNFRVYHLPKPGQIEQMTYMALENDPEAVKAOILENQCITFLGAG 359  
QY 357 FWEESDNDNMNETASQNKKKYOSRSDLSNLGFGEDVYGDAVYPGVYKSAIGCTSTRGF 416  
Db 360 LFDNDGDNDLTLACTEGRGWRFTAQMDVYTNMAGLSRGKREG-FPGDIAAGLVSEHNRYP 418

QY 417 YRAYQAHVSSSNMAE 431  
Db 419 YRRNGEHMMAETVAE 433

## RESULT 8

A65031  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escheri  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: A65031; S49292  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.: Rose, D.J.; Mau, B.; Shaoh, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; M01D:97426617; PMID:9278503  
A:Accession: A65031  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <BLAT>  
A:Cross-references: GB:AE000340; GB:U00096; NID:g1788883; PIDN:AACT5591.1; PID:g17888  
A:Experimental source: strain K-12, substrain MG1655  
R:Turtin, E.; Gasser, F.; Biville, F.  
submitted to the EMBL Data Library, September 1994  
A:Description: Cloning and sequencing of an E. coli gene homologous to dioxygenase of  
A:Reference number: S49292  
A:Accession: S49292  
A:Molecule type: DNA  
A:Residues: 1-19, 'A', 21-383, 'ATAPATANCWKW', 397-398, '451, 'R', 453, 'SAATTAELALITISFOKPL  
A:Cross-references: EMBL:Z37966; NID:g550595; PIDN:CAAB6018.1; PID:g550596  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:75-123/Domain: Rieske [2Fe-2S] homology <RSK>  
F:85,87,105,108/Binding site: 2Fe-2S cluster (Cys, His, Cys, Cys, His) (covalent) #status

Query Match 30.5%; Score 735; DB 2; Length 453;  
Best Local Similarity 36.4%; Pred. No. 1.1e-50;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGDELFQHELKTIFARNMLFLTHDSLIPAPGDVYTKMGKIDEVIVSRQDGSIRAFNL 79  
Db 24 IYDPDITQLELRFIFGRCWFLAHESQIPKPGDFNTYMGEDAVVYVRQKDSIRAFNL 83  
QY 80 VCRHNGKTLVSVEANAGFCVSYHGWFSGNGELQSYFFEKDLVGEISLNKKCLGLKEVA 139  
Db 84 QCRIRARVSYADCGNTRAFTCYHNGSYCINGELIDVLEPRARYGGLCKSHMGVNEVP 143  
QY 140 RVESFHGFIYGCEDQAPRLMDYLGDAAWYLEP-M-KHSGLELVGPGRKVVIAKMKAP 198  
Db 144 CVESYKGLIFGMWDTSAPEGLRDYLDGMLDRREGTEIVGQVQKWINCMKFP 203  
QY 199 AENYVGAIVYHGMTHASSL-----RSGESTFSSLAGMALP----- 234  
Db 204 AEOFASDQYHALFSAVAOVYLGAKDDSDKRLGD-----GQTRAPWETAKDALOG 256  
QY 235 --PRGAGLOMTSKYSGMGVLMDOYSGVH--SADLVPELMAFGAKOERLKEIGDVARLY 292  
Db 257 QDGHSGGFTEKPDANWV--DGAVSYIRETYAE-----ADGRLEVALAL 303  
QY 293 RSHLNTVFPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDMEDLKRRLADSVORTA 352  
Db 304 AGHNN--IFPFLSWLNGTATLVRWHNPRGQVYVMAFCITDKAASDEVKAAPENSATRAF 361  
QY 353 GRAGFWEESDNDNMNETASQNKKKYOSRSDLSNLGFGEDVYGDAVYPGVYKSAIGERS 412  
Db 362 GPAGFLFDQDSEWNCLEQKLLKGRHARNSKLCLEMGGOKRRDDIGIT-NYIPSETA 420  
QY 413 YRGFYRAYQAHVSSSNMAEFEEHASTWHTLT 445  
Db 421 ARGMYQWADLLESSEMQEVLDTTAAYQOEVMK 453

## RESULT 9

D91054  
 bphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escherichia  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 31-Dec-2001  
 C:Accession: D91054  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: D91054  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-453 <HAV>  
 A:Cross-references: GB:BA000007; PIDN:BA036827.1; PID:G13362875; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Keywords: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
 C:Keywords: oxidoreductase

Query Match 30.5%; Score 735; DB 2; Length 453;  
 Best Local Similarity 36.4%; Pred. No. 1.1e-50;  
 Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

OY 20 IHGDELFQHELTFFARWMLFLTHDSLIPAPGDVYTAKMIDEIVYSRONDSIRAFIN 79  
 DB 24 ITTDDIYQLELERIFGRCWMLFLAHSQIPKRGDFNTYMGDAVAVVVRQKGSIKAFIN 83  
 OY 80 VCRHKGTLVSVAGNAGKFCVSHGWFSGNGELQSVFPEKDLGSELNKKCLGEYA 139  
 DB 84 QCRHMRVSVADCGNTRAFCTPHGWSYGINGLIDVLEPRAYPQGLCKSHMGLNEVP 143  
 OY 140 RVESFHGTYGCFDDEAPPLMDYLDGDAWYLEPMF-KHSGGLELVGPPEKVIYKANWKP 198  
 DB 144 CVESYKGLFEGMWDTSAPGLRDYLDGMDLRREGGTEIVGGVQKWINCNWKP 203  
 OY 199 AENFGDAVHWGTHASSL-----RSGSIFSLGNALP----- 234  
 DB 204 AEOFPADQYHALFSAHSAVOYLGAADDGSKRLG-----GQARPYWETAKDALOG 256  
 OY 235 --PEGAGLQMTSKYSGMGVLDGYSVHSAADLVPELMAFGAKOERLNKEIGDYRARIY 292  
 DB 257 QDGHSGFEFFTEKPPANVWV--DGAVSSYRETYAE-----AEQRLGEVRLRL 303  
 OY 293 RSHLCTVFPNNSMLTSCGVFVYVNPIDANTTEVWYTAIVEKDEMDLKRRLADSVQRTA 352  
 DB 304 AGHNN--TFPTLSWLNGLATLRLVWHRGPDQVEVWAFCTTDAASDEYKAAFEENSATRA 361  
 OY 353 GPAGWESDDNDNMTASONGKKYQSRDLSNLGFGEDVYGDVAVYGVGKSAIGETS 412  
 DB 362 GPAGLEDDDSNMCEIQKLKGHRARNSKLCLEMGLOGEKRKRDGIRGIT-NTYFSETA 420  
 OY 413 YRGFYRAYQAHVSSNMAEFHASSTWHTELTK 445  
 DB 421 ARGMYQRMADLLSSEMOEVLDTAAYQOEVMK 453

RESULT 10

G85898  
 bphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escherichia  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Dec-2001  
 C:Accession: G85898  
 R:Perina, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grobbeck, E.J.; Davis, N.W.; Jim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: G85898  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-453 <STO>

A:Cross-references: GB:AE005174; NID:912516944; PIDN:ANG57651.1; GSPDB:GN00145; UMGPC:  
 A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics: hcaal  
 C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
 C:Keywords: oxidoreductase

Query Match 30.5%; Score 735; DB 2; Length 453;  
 Best Local Similarity 36.4%; Pred. No. 1.1e-50;  
 Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

OY 20 IHGDELFQHELTFFARWMLFLTHDSLIPAPGDVYTAKMIDEIVYSRONDSIRAFIN 79  
 DB 24 ITTDDIYQLELERIFGRCWMLFLAHSQIPKRGDFNTYMGDAVAVVVRQKGSIKAFIN 83  
 OY 80 VCRHKGTLVSVAGNAGKFCVSHGWFSGNGELQSVFPEKDLGSELNKKCLGEYA 139  
 DB 84 QCRHMRVSVADCGNTRAFCTPHGWSYGINGLIDVLEPRAYPQGLCKSHMGLNEVP 143  
 OY 140 RVESFHGTYGCFDDEAPPLMDYLDGDAWYLEPMF-KHSGGLELVGPPEKVIYKANWKP 198  
 DB 144 CVESYKGLFEGMWDTSAPGLRDYLDGMDLRREGGTEIVGGVQKWINCNWKP 203  
 OY 199 AENFGDAVHWGTHASSL-----RSGSIFSLGNALP----- 234  
 DB 204 AEOFPADQYHALFSAHSAVOYLGAADDGSKRLG-----GQARPYWETAKDALOG 256  
 OY 235 --PEGAGLQMTSKYSGMGVLDGYSVHSAADLVPELMAFGAKOERLNKEIGDYRARIY 292  
 DB 257 QDGHSGFEFFTEKPPANVWV--DGAVSSYRETYAE-----AEQRLGEVRLRL 303  
 OY 293 RSHLCTVFPNNSMLTSCGVFVYVNPIDANTTEVWYTAIVEKDEMDLKRRLADSVQRTA 352  
 DB 304 AGHNN--TFPTLSWLNGLATLRLVWHRGPDQVEVWAFCTTDAASDEYKAAFEENSATRA 361  
 OY 353 GPAGWESDDNDNMTASONGKKYQSRDLSNLGFGEDVYGDVAVYGVGKSAIGETS 412  
 DB 362 GPAGLEDDDSNMCEIQKLKGHRARNSKLCLEMGLOGEKRKRDGIRGIT-NTYFSETA 420  
 OY 413 YRGFYRAYQAHVSSNMAEFHASSTWHTELTK 445  
 DB 421 ARGMYQRMADLLSSEMOEVLDTAAYQOEVMK 453

RESULT 11

T31258  
 aromatic oxygenase large chain - Sphingomonas aromaticivorans plasmid pNL1  
 C:Species: Sphingomonas aromaticivorans  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
 C:Accession: T31258  
 R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom  
 A:Reference number: Z20992  
 A:Accession: T31258  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-455 <ROK>  
 A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378399; PIDN:AD03982.1  
 C:Genetics: pNL1  
 A:Note: bphA1b  
 C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
 C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
 F:88,90,108,111/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status  
 Query Match 30.0%; Score 723.5; DB 2; Length 455;  
 Best Local Similarity 36.1%; Pred. No. 9.2e-50;  
 Matches 153; Conservative 83; Mismatches 169; Indels 19; Gaps 8;

OY 2 NYNNKIYSEGLQKHLIHGDELFQHELTFFARWMLFLTHDSLIPAPGDVYTAKMGI 61  
 DB 11 DYSRYMDLKEGMLDR--IFSDADIYEELVRIFFARSWLFLVAHESQIYSSGDFLTTHNGE 68











PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
XX -  
PS Claim 9, Page 128-129, 151pp. English.

The present invention describes naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2,3,4-tetrahydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydro-1,2-dihydronaphanthrene or 3,4-dihydro-3,4-dihydronaphanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzo(cyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthrene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

**SQ Sequence 449 AA;**

Query Match	100.0%	Score 2408;	DB 21;	Length 449;
Best Local Similarity	100.0%;	Pred. No. 1,4e-225;		
Matches 449; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MNYNKKLLVSSSGISQXKHLHGDDELFQHEHKTTFARNMLFLTHDLSILPAPGDIYIAKKMG	60
Dd	1	MNYNKKLLVSSSGISQXKHLHGDDELFQHEHKTTFARNMLFLTHDLSILPAPGDIYIAKKMG	60
QY	61	IDEVIVSKONGSGIRAEFLNVCRRHGRKFLVSEAGNAGKFGVCSYHGMOFGSNGELQSVPE	120
Dd	61	IDEVIVSKONGSGIRAEFLNVCRRHGRKFLVSEAGNAGKFGVCSYHGMOFGSNGELQSVPE	120
QY	121	KDLVGEISLNNKCLGLKEVARYESFHGFTTTCGFDQDAPPLMDYIGDAAMYIEBPFKHSGL	180
Dd	121	KDLVGEISLNNKCLGLKEVARYESFHGFTTTCGFDQDAPPLMDYIGDAAMYIEBPFKHSGL	180
QY	181	ELVBPBGVYVIAKMWKPAEAEFNVDAHYHGWTHASSLRSGESIFSSLAGNAALPPRGAGL	240
Dd	181	ELVBPBGVYVIAKMWKPAEAEFNVDAHYHGWTHASSLRSGESIFSSLAGNAALPPRGAGL	240
QY	241	QMTSKYSGSGMVLWDGYSGVHSADVLBELMAFGGAKQERLNKEIGDVRARIYRSHLNCYV	300
Dd	241	QMTSKYSGSGMVLWDGYSGVHSADVLBELMAFGGAKQERLNKEIGDVRARIYRSHLNCYV	300
QY	301	FPNNSMLTCSGVFVWNPIDANTTEVWTYALVEKDWFEDLKRRLADSVQRTAAPPAFWES	360
Dd	301	FPNNSMLTCSGVFVWNPIDANTTEVWTYALVEKDWFEDLKRRLADSVQRTAAPPAFWES	360
QY	361	DDDNMMETASONGKXKYSRPSDLSNLGPFEDYVGDVYVPGVYKSAIGETSYRGFYRAY	420
Dd	361	DDDNMMETASONGKXKYSRPSDLSNLGPFEDYVGDVYVPGVYKSAIGETSYRGFYRAY	420
QY	421	QAHVSSSSNMAEEFHASSTWHTELTKYTD	449
Dd	421	QAHVSSSSNMAEEFHASSTWHTELTKYTD	449

RESULT 2	
AAB12565	
ID	AAB12565 standard; Protein; 449 AA.
XX	
AC	AAB12565;

XX	09-NOV-2000	(first entry)
DT		

DE Naphthalene dioxygenase mutant F352V protein sequence SEQ ID NO:2.

KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KW polymer; resin; pharmaceutical; rubber industry; bioremediation.

05	<i>Pseudomonas</i> sp.
05	Synthetic.

PN W02000037480-A1.

PD 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA ( IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI; 2000-452174/39.

DR N-PSDB; AAA65339.

PT Novel naphthalene dioxygenase mutant having a specific amino acid

pharmaceutical or rubber industry and for carrying out bioremediation

XX:

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, diphenyl, phenanthrene, and indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphenanthrene or 3,4-dihydroxy-3,4-dihydronaphenanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, acenaphthothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

SQ Sequence 449 AA;

Query Match	99.8%;	Score 2404;	DB 21;	Length 449;
Best Local Similarity	99.8%;	Pred. No. 3.4e-225;		
Matches 448; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	1	MNNKILVSESGLSOKHILHODEELFOHEKTIIFARNWLELTHDSLLPAPDGYVAKMG	60
Db	1	MNNKILVSESGLSOKHILHODEELFOHEKTIIFARNWLELTHDSLLPAPDGYVAKMG	60
QY	61	IDEIVYSRQNDISIRAFILNVRHGRKTLVSEAGNAKGFVCSYHMGSGNSGELQSNPFE	120
Db	61	IDEIVYSRQNDISIRAFILNVRHGRKTLVSEAGNAKGFVCSYHMGSGNSGELQSNPFE	120
QY	121	KDLVGSGLNKKCLGLKEVARVESFHFGLYGCFOQOEAFLMDLYLGDAAVLEPMKHSGL	180

Db 121 KDLGSESLNKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180  
QY 181 ELVGPPGKVYIKANKMKAPAEFNGDGYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
Db 181 ELVGPPGKVYIKANKMKAPAEFNGDGYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
Db 241 QMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGCVKWNPPIDANTTEWYTAIYEKMDPEDLKRLADSVORTGPAQFWES 360  
Db 301 FPNNSMLTCSGCVKWNPPIDANTTEWYTAIYEKMDPEDLKRLADSVORTGPAQFWES 360  
QY 361 DDNDNMETASONGKKYQSRDLSLNLGFGEDYGDYAVPGVYKSAIGETSYRGFYRAY 420  
Db 361 DDNDNMETASONGKKYQSRDLSLNLGFGEDYGDYAVPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFHASTWHTTELTKTTDR 449  
Db 421 QAHVSSSNMAEFHASTWHTTELTKTTDR 449  
RESULT 3  
AAB12567  
ID AAB12566 standard; Protein: 449 AA.  
XX  
AC AAB12566;  
DT 09-NOV-2000 (first entry)  
XX  
DE NDO related complex alpha subunit protein sequence SEQ ID NO:14.  
XX  
KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR MPI: 2000-452174/39.  
DR N-PSDB; AAA65340.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Claim 13; Page 99-100; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which

CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 449 AA;

Query Match 99.8%; Score 2404; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 3,4e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFTHHDSLIPAPGYVAKMG 60  
Db 1 MNYNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFTHHDSLIPAPGYVAKMG 60  
QY 61 IDEVIVSRONDGSIKRAFLVNCRRHKGTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
Db 61 IDEVIVSRONDGSIKRAFLVNCRRHKGTLVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLGSESLNKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180  
Db 121 KDLGSESLNKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180  
QY 181 ELVGPPGKVYIKANKMKAPAEFNGDGYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
Db 181 ELVGPPGKVYIKANKMKAPAEFNGDGYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
Db 241 QMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGCVKWNPPIDANTTEWYTAIYEKMDPEDLKRLADSVORTGPAQFWES 360  
Db 301 FPNNSMLTCSGCVKWNPPIDANTTEWYTAIYEKMDPEDLKRLADSVORTGPAQFWES 360  
QY 361 DDNDNMETASONGKKYQSRDLSLNLGFGEDYGDYAVPGVYKSAIGETSYRGFYRAY 420  
Db 361 DDNDNMETASONGKKYQSRDLSLNLGFGEDYGDYAVPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFHASTWHTTELTKTTDR 449  
Db 421 QAHVSSSNMAEFHASTWHTTELTKTTDR 449  
RESULT 4  
AAB12567  
ID AAB12567 standard; Protein: 449 AA.  
XX  
AC AAB12567;  
DT 09-NOV-2000 (first entry)  
XX  
DE NDO related complex alpha subunit protein sequence SEQ ID NO:15.  
XX  
KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.

XX 26-OCT-1998; 98US-0105575.  
PR  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR MPI: 2000-452174/39.  
DR N-PSDB: AAA65341.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
XX  
PS Claim 13; Page 100-102; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SO  
Query Match 99.8%; Score 2404; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 3.4e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNYNKKILIVSSGSLQKHLHGDELFQHELTTFARMNMLFTHDSLIPAPEDVYATAMG 60  
DB 1 MNYNKKILIVSSGSLQKHLHGDELFQHELTTFARMNMLFTHDSLIPAPEDVYATAMG 60  
QY 61 IDEVIVSRONDGSIIRAFNLNCRHKGKTLVSVEAGNAKGFVCSYHGWGFGSNGLQSVPE 120  
DB 61 IDEVIVSRONDGSIIRAFNLNCRHKGKTLVSVEAGNAKGFVCSYHGWGFGSNGLQSVPE 120  
QY 121 KDLVGSLENKKCLGLKEVARYESTHGFYCGFDEAPPLMDYLGDAAVLEPMFHHSGGL 180  
DB 121 KDLVGSLENKKCLGLKEVARYESTHGFYCGFDEAPPLMDYLGDAAVLEPMFHHSGGL 180  
QY 121 KDLVGSLENKKCLGLKEVARYESTHGFYCGFDEAPPLMDYLGDAAVLEPMFHHSGGL 180  
DB 121 KDLVGSLENKKCLGLKEVARYESTHGFYCGFDEAPPLMDYLGDAAVLEPMFHHSGGL 180  
QY 181 ELVGPPEKVIKAMWKAPEAFNVGDAYHVGMTTHASLSRGSSTFSSLAGNALPREGAGL 240  
DB 181 ELVGPPEKVIKAMWKAPEAFNVGDAYHVGMTTHASLSRGSSTFSSLAGNALPREGAGL 240  
QY 241 QMTSKYSGSGVGLMDGYSGVSHADVLPELMAFGAKQERLKEIGDVARIYRSHLNCYV 300  
DB 241 QMTSKYSGSGVGLMDGYSGVSHADVLPELMAFGAKQERLKEIGDVARIYRSHLNCYV 300  
QY 301 FPNNSMLTSCSVFVYVWNPIDANTTEWWTYAIVEKMPEDLKRRLADSVQRTAGPAFEMES 360  
DB 301 FPNNSMLTSCSVFVYVWNPIDANTTEWWTYAIVEKMPEDLKRRLADSVQRTAGPAFEMES 360  
QY 361 DDNDNMETASONGKKYKSRDLSLNLGEGEDVYGDAAVPCVWGSATGEMSYRGFYRAY 420  
DB 361 DDNDNMETASONGKKYKSRDLSLNLGEGEDVYGDAAVPCVWGSATGEMSYRGFYRAY 420

QY 421 QAHVSSNMAREFHASTWHTLFTKTD 449  
DB 421 QAHVSSNMAREFHASTWHTLFTKTD 449  
RESULT 5  
AAB12578  
ID AAB12578 standard; Protein: 449 AA.  
XX  
XX AAB12578;  
XX  
XX 09-NOV-2000 (first entry)  
XX  
XX Naphthalene dioxygenase mutant F352G protein sequence SPQ ID NO:32.  
XX  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
XX Pseudomonas sp.  
OS Synthetic.  
XX  
XX WO200037480-A1.  
XX  
XX 29-JUN-2000.  
XX  
XX 26-OCT-1999; 99WO-US25079.  
PF 26-OCT-1998; 98US-0105575.  
XX  
XX 26-OCT-1998; 98US-0105575.  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX MPI: 2000-452174/39.  
DR N-PSDB: AAA65352.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
XX Claim 9; Page 126-127; 151pp; English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SO  
Query Match 99.8%; Score 2404; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 3.4e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MNYNKKILVSSGSLQKHLHGDEELQHELEKTIIPARNMLPLTHDLSIPAPGDTVAKMG	60
Db	1	MNYNKKILVSSGSLQKHLHGDEELQHELEKTIIPARNMLPLTHDLSIPAPGDTVAKMG	60
QY	61	IDEVTVSRQNDGSRAPFLNVRHRGKTLVSYEAGNAGFCVSHGMFGSNGELQSPFE	120
Db	61	IDEVTVSRQNDGSRAPFLNVRHRGKTLVSYEAGNAGFCVSHGMFGSNGELQSPFE	120
QY	121	KDLVGSJLNKKCJLCLKEVARVSEPHGEIYGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL	180
Db	121	KDLVGSJLNKKCJLCLKEVARVSEPHGEIYGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL	180
QY	181	ELVGPFGKVVITKANMKAPAEENFVGDAYHVCGTHASSLRSGESITFSSLAGNAALPREGAGL	240
Db	181	ELVGPFGKVVITKANMKAPAEENFVGDAYHVCGTHASSLRSGESITFSSLAGNAALPREGAGL	240
QY	241	QMTSKYSGGMVLMDGVSNGVSHADYVPELMAFGAKOERLNKELTGDVRAITYSHLNCTV	300
Db	241	QMTSKYSGGMVLMDGVSNGVSHADYVPELMAFGAKOERLNKELTGDVRAITYSHLNCTV	300
QY	301	FPNNSMLTLCGCVFVKNPPIDANTLEWVTYIAIVEKDMPDEDLKRLADSVCQRTAPAGHWES	360
Db	301	FPNNSMLTLCGCVFVKNPPIDANTLEWVTYIAIVEKDMPDEDLKRLADSVCQRTAPAGHWES	360
QY	361	DDNDNMETASONGKKKYOSRSDLSNLGFGEDGYGDAVYGVGVGKSAIGETSYRGFPRAY	420
Db	361	DDNDNMETASONGKKKYOSRSDLSNLGFGEDGYGDAVYGVGVGKSAIGETSYRGFPRAY	420

OY	421	QAHVSSSNMAFEFHAASSTWHTLTKTDR	449
Db	421	QAHVSSSNMAFEFHAASSTWHTLTKTDR	449
RESULT	6		
ID	AA12580		
XX	AA12580	standard; Protein; 449 AA.	
AC	AA12580;		
XX			
DT	09-NOV-2000	(first entry)	
XX			
DE	Naphthalene dioxygenase mutant F352T	protein sequence SEQ ID NO:34.	
KW	Pseudomonas sp. strain NCIB 9816-4;	naphthalene dioxygenase; NDO;	
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
XX			
OS	Pseudomonas sp.		
OS	Synthetic.		
XX			
PN	WO200037480-A1.		
XX			
PD	29-JUN-2000.		
XX			
PF	26-OCT-1999;	99MO-US25079.	
XX			
PR	26-OCT-1998;	98US-0105575.	
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
XX			
PI	Parales R, Gibson D, Resnick S, Lee K;		
XX			
DR	WPI: 2000-452174/39.		
DR	N-PSDB; AAA65354.		
PT	Novel naphthalene dioxygenase mutant having a specific amino acid		
PT	substitution for preparing chiral diols for use in the polymer, resin		
PT	pharmaceutical or rubber industry and for carrying out bioremediation		
XX			
XX			
PS	Claim 9; Page 129-131; 151pp; English.		

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (1) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydrodiol, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzo(c)lone(1t)-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothiophene, 9,10-dihydronaphthene, or 9,10-dihydronaphanthrene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Query Match	99.8%	Score 2404	DB 21	Length 449
Best Local Similarity	99.8%	Pred. No. 3.4e-225		
Matches 448; Conservative	0	Mismatches 1	Indels 0	Gaps 0

[illegible]

KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol,
KX	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
OS	Pseudomonas sp.
OS	Synthetic.
XX	
PN	WO200037480-A1.
XX	
PD	29-JUN-2000.
XX	
PF	26-OCT-1999; 99MO-US25079.
XX	
PR	26-OCT-1998; 98US-0105575.
XX	
PA	(IOWA ) UNIV IOWA RES FOUND.
XX	
PI	Parales R, Gibson D, Resnick S, Lee K;
XX	
DR	WPI: 2000-452174/39.
DR	N-PSDB; AAA65355.

Novel naphthalene dioxygenase mutant having a specific amino acid substitution for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry and for carrying out bioremediation

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (1) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphanthrene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Query Match	99.8%	Score 2403;	DB 21;	Length 449;
Best Local Similarity	99.8%	Pred. No. 4.3e-225;		
Matches 448;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	MNNKLLVSSSGLSQKHLHGHDEELFQHHEKTIFFANNMLFLTHDSLIPRPGDYVAKMG	60
Db	1	MNYNKLIVSESSGISOQKHLHGHDEELFQHHEKTIFFANNMLFLTHDSLIPRPGDYVAKMG	60
Qy	61	IDEVYVSRONGCSIRAFILNVCRRHGKFLTVSEAGNAGKFCVSHGMGFGSGELQSYPE	120
Db	61	IDEVYVSQNDGCSIRAFILNVCRRHGKFLTVSEAGNAGKFCVSHGMGFGSGELQSYPE	120
Qy	121	KDLVGESLNKKCLGLKEVAVSEFHGEIYGCFODEAPRLMDYLGDAWYLEPFFKHSGGI	180
Db	121	KDLVGESLNKKCLGLKEVAVSEFHGEIYGCFODEAPRLMDYLGDAWYLEPFFKHSGGI	180
Qy	181	ELVPPKGVYVTKAMWKPAENFVGDAYHVGTHASSLSRSGESITFSSLAGNAALPPRAGI	240
Db	181	ELVPPKGVYVTKAMWKPAENFVGDAYHVGTHASSLSRSGESITFSSLAGNAALPPRAGI	240

Qy	241	QMTSKYGGMGVLMGYSYGVSHADLYPELPMAFGGAQKERLNKEIGDVRAIITYSHLNCIY	300
Db	241	QMTSKYGGMGVLMGYSYGVSHADLYPELPMAFGGAQKERLNKEIGDVRAIITYSHLNCIY	300
Qy	301	FPNNMLTCSGCFKFWNP>IDANTTEVWYTYAIVEKIDMEEDLKRRLADSVORTAPAGWES	360
Db	301	FPNNMLTCSGCFKFWNP>IDANTTEVWYTYAIVEKIDMEEDLKRRLADSVORTAPAGWES	360
Qy	361	DDNDMMETASONGKKYQSRDSDLNLNLCFGEDVYGDAVYPCGVGSAIIGETSYRGFYRAY	420
Db	361	DDNDMMETASONGKKYQSRDSDLNLNLCFGEDVYGDAVYPCGVGSAIIGETSYRGFYRAY	420
Qy	421	QAHVSSSWMAEPFAHASTTWHTELRKTTDR	449
Db	421	QAHVSSSWMAEPFAHASTTWHTELRKTTDR	449

RESULT 8  
AAB12582  
ID AAB12582 standard; Protein; 449 AA.

AC AAB12582;

DT 09-NOV-2000 (first entry)

DE Naphthalene dioxygenase mutant F352I protein sequence SEQ ID NO:36.

KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol,  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS	Pseudomonas sp.
OS	Synthetic.

PN WO2000037480 -

PD 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA ( IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPT; 2000-452174/39.

XX

PT substitution for preparing chiral diols for use in the polymer, resin

[illegible]

PS Claim 9; Page 133-134; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (1) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +/-)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphenanthrene or 3,4-dihydroxy-3,4-dihydronaphenanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphenanthrene a



CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA;

Query Match 99.8%; Score 2403; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 4.3e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNKILVSESGLSQKHLHGDELFQHELKTIFARNMLFLTHDSLIPAGDYVYAKMG 60  
DB 1 MNNKILVSESGLSQKHLHGDELFQHELKTIFARNMLFLTHDSLIPAGDYVYAKMG 60  
QY 61 IDEVIYSRONDGSIRAFNLVNCRRHRTLVSEGNKAGFYCSYHGMFGSNGELQSVPEE 120  
DB 61 IDEVIYSRONDGSIRAFNLVNCRRHRTLVSEGNKAGFYCSYHGMFGSNGELQSVPEE 120  
QY 121 KDLYGESLNKKCLGLEKVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPFRKHSGL 180  
DB 121 KDLYGESLNKKCLGLEKVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPFRKHSGL 180  
QY 181 ELVGPBGKVIYKANKMKAPAEVNDAYHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
DB 181 ELVGPBGKVIYKANKMKAPAEVNDAYHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVHSHADLPELMARFGAEOERLNKEIGVRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMDSYGVHSHADLPELMARFGAEOERLNKEIGVRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKMPEDLKRRLADSVORTGAPAGFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKMPEDLKRRLADSVORTGAPAGFWES 360  
QY 361 DDNDNMETASONGKRYQSRDLSNLGEGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASONGKRYQSRDLSNLGEGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNNAAEFPHASSTWTELTCTTDR 449  
DB 421 QAHVSSNNAAEFPHASSTWTELTCTTDR 449

RESULT 9

AA12577  
ID AAB12577 standard; Protein; 449 AA.

AC AAB12577;

DT 09-NOV-2000 (first entry)

DE Naphthalene dioxygenase alpha subunit protein sequence SEQ ID NO:26.

XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
XX polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS Pseudomonas sp.

PN WO200037480-A1.

PD 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB: AAA65351.

PT Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,

PT pharmaceutical or rubber industry and for carrying out bioremediation

XX Disclosure; page 118-120; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzochiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC naphthalene dioxygenase (NDO) from the Pseudomonas sp. strain NCIB  
CC 9816-4, which is used in the exemplification of the present invention.

XX Sequence 449 AA;

Query Match 99.8%; Score 2402; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 5.4e-225;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNKILVSESGLSQKHLHGDELFQHELKTIFARNMLFLTHDSLIPAGDYVYAKMG 60  
DB 1 MNNKILVSESGLSQKHLHGDELFQHELKTIFARNMLFLTHDSLIPAGDYVYAKMG 60  
QY 61 IDEVIYSRONDGSIRAFNLVNCRRHRTLVSEGNKAGFYCSYHGMFGSNGELQSVPEE 120  
DB 61 IDEVIYSRONDGSIRAFNLVNCRRHRTLVSEGNKAGFYCSYHGMFGSNGELQSVPEE 120  
QY 121 KDLYGESLNKKCLGLEKVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPFRKHSGL 180  
DB 121 KDLYGESLNKKCLGLEKVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPFRKHSGL 180  
QY 181 ELVGPBGKVIYKANKMKAPAEVNDAYHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
DB 181 ELVGPBGKVIYKANKMKAPAEVNDAYHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVHSHADLPELMARFGAEOERLNKEIGVRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMDSYGVHSHADLPELMARFGAEOERLNKEIGVRARIYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKMPEDLKRRLADSVORTGAPAGFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVWYAIYEKMPEDLKRRLADSVORTGAPAGFWES 360  
QY 361 DDNDNMETASONGKRYQSRDLSNLGEGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASONGKRYQSRDLSNLGEGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNNAAEFPHASSTWTELTCTTDR 449  
DB 421 QAHVSSNNAAEFPHASSTWTELTCTTDR 449

RESULT 10

ID	AAB12583
AC	AAB12583 standard; Protein; 449 AA.
XX	
XX	AAB12583;
DT	09-NOV-2000 (first entry)
XX	
DE	Naphthalene dioxygenase mutant F352W protein sequence SEQ ID NO:58.
KX	
RW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
XX	polymer; resin; pharmaceutical; rubber industry; bioremediation.
OS	Pseudomonas sp.
OS	Synthetic.
PN	WO200037480-A1.
PD	
XX	29-JUN-2000.
PF	
XX	26-OCT-1999; 99MO-US25079.
PR	
XX	26-OCT-1998; 98US-0105575.
PA	(IOWA ) UNITV IOWA RES FOUND.
PI	
DR	Parales R, Gibson D, Resnick S, Lee K;
XX	WPI; 2000-452174/39.
PT	N-PsDB; AAA65376.
PT	
PT	Novel naphthalene dioxygenase mutant having a specific amino acid
PT	substitution for preparing chiral diols for use in the polymer, resin,
PT	pharmaceutical or rubber industry and for carrying out bioremediation
PS	-
XX	
XX	Example 7; Page 142-144; 151pp; English.
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO
CC	related complex (I) comprising several polypeptides which contain an
CC	alpha subunit that contains substituted amino acids at specific
CC	positions. The polypeptides and host cells are useful for preparing
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-
CC	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-
CC	dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which
CC	involves contacting them with phenanthrene. The polypeptides and the
CC	host cells are also used in bioremediation in which they oxidise an
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,
CC	benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,
CC	acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
CC	dienanthrophenone, 9,10-dihydroanthracene, or 9,10-dihydronaphthrene a
CC	corresponding dihydro dihydroxy compound. The polypeptides and the host
CC	cells are useful for preparing chiral diols for use in the polymer,
CC	resin, pharmaceutical or rubber industry. The present sequence represents
CC	a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
CC	exemplification of the present invention.
XX	
XX	Sequence 449 AA:
SO	
Query Match	99.7%; Score 2401; DB 21; Length 449;
Best Local Similarity	99.8%; Pred. No. 6; 7e-225;
Matches 448; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 MNYNKKILVSESGLSOKHLIHGDELDFQHELKTFIRARNLFLTHDSLIPAGDYVTAKWG 60
Db	1 MNYNKKILVSESGLSOKHLIHGDELDFQHELKTFIRARNLFLTHDSLIPAGDYVTAKWG 60
OY	61 IDEIVYSQNDSIFAFLNVCRRRGKITLVSVENGAAGKGVCSYHGMGFSSNELGSVPPE 120

Db	61	IDVIVSRQNDGSIIRAFLLNCRHNGKTLVSEVGNKKNGFCVSHGNGFSGNELDQSVFE	120
Qy	121	KDLVGESENKKCLGLKEVAREYSEFHGFIYCCFDQDEAPPLMDYLGDAWYLEPMFRHSGGL	180
Db	121	KDLVGESENKKCLGLKEVAREYSEFHGFIYCCFDQDEAPPLMDYLGDAWYLEPMFRHSGGL	180
Qy	181	ELVGPGRGVYIKAWKPAEAEFVGDAAHVGWTHASSLRGSESTFSSLAGNALPREGAGL	240
Db	181	ELVGPGRGVYIKAWKPAEAEFVGDAAHVGWTHASSLRGSESTFSSLAGNALPREGAGL	240
Qy	241	QMTSKYSGMGVLWDGYSGVASADLVPELMAFGAKOERLNKEIGVRRARIYRSHLNTCV	300
Db	241	QMTSKYSGMGVLWDGYSGVASADLVPELMAFGAKOERLNKEIGVRRARIYRSHLNTCV	300
Qy	301	FPNNSMLTCSGVFVYVWNPIDANTTEVWTVAIVKEKMDPIKRRRLADSVQRTAGPAGFEMES	360
Db	301	FPNNSMLTCSGVFVYVWNPIDANTTEVWTVAIVKEKMDPIKRRRLADSVQRTAGPAGFEMES	360
Qy	361	DDNDMETFASONGKRYOSRSDLSLNGFSEDDYVGAIVPGVYKSAIGETSRGFAY	420
Db	361	DDNDMETFASONGKRYOSRSDLSLNGFSEDDYVGAIVPGVYKSAIGETSRGFAY	420
Qy	421	QAHVSSSNMAEFEEHASSSTWHTELTKTTDR	449
Db	421	QAHVSSSNMAEFEEHASSSTWHTELTKTTDR	449
RESULT 11			
AA12568			
ID	AA12568	standard; Protein; 449 AA.	
AC	AA12568;		
DT	09-NOV-2000	(first entry)	
DE	NDO	related complex alpha subunit protein sequence SEQ ID NO:16.	
KM	Pseudomonas sp. strain NCIB 9816-4;	naphthalene dioxygenase; NDO;	
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
KV	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
OS	Pseudomonas sp.		
OS	Synthetic.		
PN	MO200037480-A1.		
PD	29-JUN-2000.		
PF	26-OCT-1999;	99WO-US25079.	
PR	26-OCT-1998;	98US-0105575.	
PA	(IOWA ) UNIV IOWA RES FOUND.		
PI	Parales R, Gibson D, Resnick S, Lee K;		
DR	WPI: 2000-452174/39.		
DR	N-PSDB: AAA65342.		
PT	Novel naphthalene dioxygenase mutant having a specific amino acid		
PT	substitution for preparing chiral diols for use in the polymer resin,		
PT	pharmaceutical or rubber industry and for carrying out bioremediation		
PS	Claim 13; Page 102-103; 151pp; English.		
XX	The present invention describes a naphthalene dioxygenase (NDO) or NDO		
XX	related complex (1) comprising several polypeptides which contain an		
XX	alpha subunit that contains substituted amino acids at specific		
XX	positions. The polypeptides and host cells are useful for preparing		
XX	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-		
XX	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-		
XX	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which		

CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.6%; Score 2399; DB 21; Length 449;  
 Best Local Similarity 99.6%; Pred. No. 1.1e-224;  
 Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MNVNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60  
 DB 1 MNVNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60  
 QY 61 IDEVIVSRONDGSIKFLVNCRRHRTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 DB 61 IDEVIVSRONDGSIKFLVNCRRHRTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPGRKVIYANKKAPENFVGDAYHYGWTTHASSLRGSESTFSSLAGNALPREGAQL 240  
 DB 181 ELVGPGRKVIYANKKAPENFVGDAYHYGWTTHASSLRGSESTFSSLAGNALPREGAQL 240  
 QY 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKQOEELNKEIDGVRRIRSHLNCYV 300  
 DB 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKQOEELNKEIDGVRRIRSHLNCYV 300  
 QY 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKDMPEDLKRLADSVORTAGPAGFWES 360  
 DB 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKDMPEDLKRLADSVORTAGPAGFWES 360  
 QY 361 DDNDNNKETAQNGKRYQSDLSLNLGEGEDYVGDVYPVYGSAIETSYRGYRAY 420  
 DB 361 DDNDNNKETAQNGKRYQSDLSLNLGEGEDYVGDVYPVYGSAIETSYRGYRAY 420  
 QY 421 QAHVSSSNMAEFEBHASTWTELTCTTDR 449  
 DB 421 QAHVSSSNMAEFEBHASTWTELTCTTDR 449

RESULT 12  
 AAB12584  
 ID AAB12584 standard: Protein; 448 AA.  
 XX AAB12584;  
 XX  
 DT 09-NOV-2000 (first entry)  
 DE Naphthalene dioxygenase mutant F352Y protein sequence SEQ ID NO:59.  
 XX  
 KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 XX Pseudomonas sp.  
 OS Synthetic.  
 OS  
 PN WO200037480-A1.

XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 PF 26-OCT-1998; 98US-0105575.  
 PR (IOWA ) UNIV IOWA RES FOUND.  
 PA  
 PI Parales R, Gibson D, Resnick S, Lee K;  
 DR WPI: 2000-452174/39.  
 DR N-PSDB: AAA65377.  
 PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 -

Example 7; Page 144-145; 151pp: English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

SO Sequence 448 AA:

Query Match 99.5%; Score 2397; DB 21; Length 448;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-224;  
 Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60  
 DB 1 MNVNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60  
 QY 61 IDEVIVSRONDGSIKFLVNCRRHRTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 DB 61 IDEVIVSRONDGSIKFLVNCRRHRTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
 QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPGRKVIYANKKAPENFVGDAYHYGWTTHASSLRGSESTFSSLAGNALPREGAQL 240  
 DB 181 ELVGPGRKVIYANKKAPENFVGDAYHYGWTTHASSLRGSESTFSSLAGNALPREGAQL 240  
 QY 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKQOEELNKEIDGVRRIRSHLNCYV 300  
 DB 241 QMTSKYSGMGVLMDSYGVSHSADLVPELMARFGAKQOEELNKEIDGVRRIRSHLNCYV 300  
 QY 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKDMPEDLKRLADSVORTAGPAGFWES 360  
 DB 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKDMPEDLKRLADSVORTAGPAGFWES 360

QY	361	DDNDNMTASQNKRRKQSRSDLLSNLGFGEVDYGCDAVYECVVKSAIGETSYNGFRAY	420
Db	361	DDNDNMTASQNKRRKQSRSDLLSNLGFGEVDYGCDAVYECVVKSAIGETSYNGFRAY	420
OY	421	QAHVSSSNNAEFFHASTWHTELTKTTD	448
Db	421	QAHVSSSNNAEFFHASTWHTELTKTTD	448
RESULT 13			
ID	AAB12569		
AA	AAB12569	standard; Protein; 449 AA.	
XX			
AC	AAB12569:		
XX			
DT	09-NOV-2000	(first entry)	
XX			
DE	NDO related complex alpha subunit protein sequence SEQ ID NO:17.		
XX			
KM	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase: NDO;		
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
RW	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
XX			
OS	Pseudomonas sp.		
XX	Synthetic.		
PN	WO200037480-A1.		
XX			
PD	29-JUN-2000.		
XX			
PF	26-OCT-1999; 99WO-US25079.		
XX			
PR	26-OCT-1998; 98US-0105575.		
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
PI	Parales R, Gibson D, Resnick S, Lee K;		
PT	WPI. 2000-452174/39.		
DR	N-PSDB; AAA65343.		
XX			
PT	Novel naphthalene dioxygenase mutant having a specific amino acid		
PT	substitution for preparing chiral diols for use in the polymer, resin,		
PT	pharmaceutical or rubber industry and for carrying out bioremediation		
PT	-		
XX			
PS	Claim 13; Page 104-105; 151pp; English.		
XX			
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO		
CC	related complex (I) comprising several polypeptides which contain an		
CC	alpha subunit that contains substituted amino acids at specific		
CC	positions. The polypeptides and host cells are useful for preparing		
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-		
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-		
CC	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which		
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,		
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and		
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-		
CC	dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which		
CC	involves contacting them with phenanthrene. The polypeptides and the		
CC	host cells are also used in bioremediation in which they oxidise an		
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,		
CC	benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,		
CC	acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,		
CC	dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphanthrene a		
CC	corresponding dihydro dihydroxy compound. The polypeptides and the host		
CC	cells are useful for preparing chiral diols for use in the polymer,		
CC	resin, pharmaceutical or rubber industry. The present sequence represents		
CC	a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the		
CC	exemplification of the present invention.		
XX			
Sequence	449 AA:		
XX			

	Query Match	99.4%	Score 2394;	DB 21;	Length 449;
	Best Local Similarity	99.3%	Pred. No. 3.2e-224;		
	Matches 446; Conservative	1;	Mismatches 2;	Indels 0;	Gaps
QY	1 MMYNNKILVSESGLSOKLHIGDEDFPHELEKTTFARNWLFPTHDSLIPAPGVYTKMG	60			
Db	1 MMYNNKILVSESGLSQKHLHGDELFPHELKTFARNWLFPTHDSLIPAPGVYTKMG	60			
QY	61 IDEIVSRQNDGSIRAFPLNVCRRHKCTLVSVEAGNAKGFCVCSYHGMFGSNGELQSVPFE	120			
Db	61 IDEIVSRQNDGSIIRAFPLNVCRRHKCTLVSVEAGNAKGFCVCSYHGMFGSNGELQSVPFE	120			
QY	121 KDLYESLNKKCKLCLEVARVESFHGFYYCGFDDEAPPFLMDYLGDAAWYLEPMFKHSGL	180			
Db	121 KDLYESLNKKCKLCLEVARVESFHGFYYCGFDDEAPPFLMDYLGDAAWYLEPMFKHSGL	180			
QY	181 ELVGPFGKVVIKAMKAFAENFVGDAIVGWTHASSLRSGESTIFSSLAGNALPPREGAGL	240			
Db	181 ELVGPFGKVVIIKAMKAFAENFVGDAIVGWTHASSLRSGESTIFSSLAGNALPPREGAGL	240			
QY	241 QMTSKYSGSMGYLVMDGYSVGHASDLYPELMAFGAKOERLKEIGDVARRIYRSHLNCYV	300			
Db	241 QMTSKYSGSMGYLVMDGYSVGHASDLYPELMAFGAKOERLKEIGDVARRIYRSHLNCYV	300			
QY	301 FPNNSMLTSCGYFFKYWNVIDANTTEWTYIAIVEKDMPEDLKRRRLADSVQRTAQPAGES	360			
Db	301 FPNNSMLTSCGYFFKYWNVIDANTTEWTYIAIVEKDMPEDLKRRRLADSVQRTAQPAGES	360			
QY	361 DDNDNMETASONGKKYGQRSDDLNLGFGEDVDYGDVAIVPGVVKSAIGETSYRGFYRAY	420			
Db	361 DDNDNMETASONGKKYGQRSDDLNLGFGEDVDYGDVAIVPGVVKSAIGETSYRGFYRAY	420			
QY	421 QAHVSSSNWAEFEHHASTFWHTELKTTDR	449			
Db	421 RAHVSSSNWAEFEHHASTFWHTELKTTDR	449			
RESULT 14					
AAB12570	.				
ID	AAB12570 standard; Protein; 449 AA.				
XX					
AC	AAB12570;				
XX					
DT	09-NOV-2000 (first entry)				
XX					
DE	NDO related complex alpha subunit protein sequence SEQ ID NO.18.				
XX					
KM	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;				
KM	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;				
XX	polymer; resin; pharmaceutical; rubber industry; bioremediation.				
OS	Pseudomonas sp.				
XX	Synthetic.				
PN	WO200037480-A1.				
PD	29-JUN-2000.				
XX					
PF	26-OCT-1999; 99WO-US25079.	*			
XX					
PR	26-OCT-1998; 98US-0105575.				
XX					
PA	(IOWA ) UNIV IOWA RES FOUND.				
XX					
PI	Parales R, Gibson D, Resnick S, Lee K;				
XX					
DR	WPI: 2000-452174/39.				
DR	N-PSDB; AAA65344.				
XX					
PT	Novel naphthalene dioxygenase mutant having a specific amino acid				
PT	substitution for preparing chiral diols for use in the polymer, resin				
PT	pharmaceutical or rubber industry and for carrying out bioremediation				

Claim 13; Page 105-107; 151pp; English.

PT -  
XX  
PS Claim 13; Page 105-107; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (1) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-diphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphenanthrene or 3,4-dihydroxy-3,4-dihydronaphenanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothienophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphenanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

Sequence 449 AA;

Query Match	97.48;	Score 2346;	DB 21;	Length 449;
Best Local Similarity	96.78;	Pred. No. 1.5e-219;		
Matches 434;	Conservative 10;	Mismatches 5;	Indels 0;	Gaps 0

QY	1	MNYNNKLIVSSGSGJQHKLHLHGBELRQHEHKLRTFARNWMLFTHDSLTPAPGDVYAKMG	60
Dp	1	MNYNNKLIVSSGSLTQKHLHLHGBELRQHEHKLRTFARNWMLFTHDSLTPBGDYAKMG	60
QY	61	IDEIVYSRÖNDGSTRAFJLWNCNRHGRKTLVSVENAGMAKGFVCSYHGMFGSNGELÖSVFE	120
Dp	61	IDEIVYISROSQSGIRAFJLWNCNRHGRKTLVNAEAGNAGFVCSYHGMFGSNGELÖSVFE	120
QY	121	KDLGSELNKKCGJLKEVAVRESHGFTYSCFOEAPJLMDYLGDAAWYLEPMPKHSGL	180
Dp	121	KELGSELNKKCGJLKEVAVRESHGFTYSCFOEAPJLMDYLGDAAWYLEPILFKHSGL	180
QY	181	ELVGPCKVYIKAMWKAAPAEJFVGDAAHYHVTHTASSLRSEJTFSSLAGNAALPREBAGL	240
Dp	181	ELVGPCKVYIKAMWKAAPAEJFVGDAAHYHVTHTASSLRSEJTFASLAGNAALPREBAGL	240
QY	241	ÖMTSKYSGMGVLMDGYSGVSHADVLPELMAFGCAKORLNKEIGDYRARIYRSHLNTCV	300
Dp	241	ÖMTSKYSGMGVLMDGYSGVSHADVLPELMAFGSKORLNKEIGDYRARIYRSHLNTCV	300
QY	301	FPNNSMLTSGSVFVWMPIDANTTEWYTVAIVEKDMPEDLKRRLADSVOQTAPBAGWES	360
Dp	301	FPNNSMLTSGSVFVWMPIDANTTEWYTVAIVEKDMPEDLKRRLADSVOQTAPBAGWES	360
QY	361	DDNDNMTASONGKKYOSRSDLSNLNGFEDYGDAPVGYGVKSAIGETSYRGFFRAY	420
Dp	361	DDNDNMTASONGKKYOSRSDLSNLNGFCKDYGDAPVGYGVKSAIGETSYRGFFRAY	420
QY	421	QAHVSSSNMAEERHASTWHTTELTKTTDR	449
Dp	421	QAHVSSSNMAEERHASTWHTTELTKTTDR	449

## RESULT 15

AAB12571

ID AAB12571 standard; Protein; 449 AA.

AC AAB12571;

DT 09-NOV-2000 (first entry)

XX	NDO related complex alpha subunit protein sequence SEQ ID NO:19.
DE	
XX	<i>Pseudomonas</i> sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
OS	<i>Pseudomonas</i> sp.
OS	Synthetic.
XX	
XX	
Key	Location/Qualifiers
FT	Misc-difference 35
FT	/note= "unspecified"

PN WO200037480-A1.  
XX  
XX  
PD 29-JUN-2000.  
XX  
PE 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI; 2000-452174/39.  
DR N-PSDB; AAA65345.  
XX

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT -  
XX  
XS Claim 13: Page 107-109; 151pp: English.

Claim 13; Page 107-109; 151pp; English.

The present invention describes anaspathalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1*S*,2*R*)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1*S*,2*R*)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphenanthrene or 3,4-dihydroxy-3,4-dihydronaphenanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohexyl-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, acenaphthophene, 9,10-dihydroanthracene, or 9,10-dihydronaphenanthrene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Sequence 449 AA;

Query Match	96.3%	Score 2318;	DB 217;	Length 449;
Best Local Similarity	95.5%	Pred. No. 8.2e-217;		
Matches 429; Conservative	12;	Mismatches 8;	Indels 0;	Gaps 0;

QY	1	MNNKNTLVSESGLSQKHLHGDDELPOHEHKTFTFANNWLFETLHDSLTPLPAGDVTAKMG	60
	1	MNNKNTLVSESGTQXKHLHGDDELPOHEHRTIYKANNWLFETLHDSLTLPBGDVTAKMG	60
Db			
	61	IDEYIYSRQDGSSTRAFLANVCRRHKGTLVSYEAGNAKGYVCSHGWFSGSNGELQSYVFE	120
QY	61	IDEYIYSRQDGSSTRAFLANVCRRHKGTLVYAEAGNAGFYVCSHGWFSGSNGELQSYVFE	120
Db			

```
QY 121 KDLYGESLNKKCLGKEVAVESFHGFIYGCFFDOEAPPLMDYLGDAWYLEPEMKHSGGL 180
|:|||||
Db 121 KELYGESLNKKCLGKEVAVESFHGFIYGCFFDOEAPPLMDYLGDAWYLEPEMKHSGGL 180
|:|||||
QY 181 ELVGPPGKVVIKANWKAFAENFVGDAYHVGMTTHASSLSRSGESIFSSLAGNNAALPPEGAGL 240
|:|||||
Db 181 ELVGPPGKVVIKANWKAFAENFVGDAYHVGMTTHASSLSRSGESIFSSLAGNNAALPPEGAGL 240
|:|||||
QY 241 QMTSKYSGMGVLMDGYSGVSHADLVPELMAFGAKOERLNKETGDVBARARYRSHLNCV 300
|:|||||
Db 241 QMTSKYSGMGVLMDGYSGVSHADLVPELMAFGAKOERLNKETGDVBARARYRSHLNCV 300
|:|||||
QY 301 FPNNMLTCSGVFVWNPIDANTTEVWMTYAIVEKDMPEDLKRLADSVQRTAGPAGFWS 360
|:|||||
Db 301 FPNNMLTCSGVFVWNPIDANTTEVWMTYAIVEKDMPEDLKRLADSVQRTAGPAGFWS 360
|:|||||
QY 361 DDNDNMETASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVVGKSAIGETSYRGFYRAY 420
|:|||||
Db 361 DDNDNMETASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVVGKSAIGETSYRGFYRAY 420
|:|||||
QY 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449
|:|||||
Db 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449
|:|||||
```

Search completed: January 27, 2003, 08:58:48  
Job time : 30.1667 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 09:02:32 : Search time 7 Seconds

(without alignments)  
1294.310 Million cell updates/sec

Title: US-09-843-250-34

Perfect score: 2409

Sequence: 1 MNVNNKILVSESGLSQKHLI.....AEFEHASTWHTELTKTTDR 449

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCr\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/PCrUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	735	30.5	453	10 US-09-815-242-10253	Sequence 10253, A
2	375.5	15.6	480	10 US-09-815-242-11692	Sequence 11692, A
3	375	15.6	490	9 US-09-738-626-6140	Sequence 6140, Ap
4	336.5	14.0	424	10 US-09-815-242-5097	Sequence 5097, Ap
5	185	7.7	35	10 US-09-776-490-19	Sequence 19, Appl
6	185	7.7	35	10 US-09-776-491-19	Sequence 19, Appl
7	174	7.2	35	10 US-09-776-490-20	Sequence 20, Appl
8	174	7.2	35	10 US-09-776-491-20	Sequence 20, Appl
9	168	7.0	35	10 US-09-776-490-21	Sequence 21, Appl
10	168	7.0	35	10 US-09-776-491-21	Sequence 21, Appl
11	118.5	4.9	354	1 US-08-976-063C-4	Sequence 4, Appl1
12	110	4.6	35	10 US-09-776-490-12	Sequence 12, Appl
13	110	4.6	35	10 US-09-776-491-12	Sequence 12, Appl
14	109	4.5	35	10 US-09-776-490-14	Sequence 14, Appl
15	109	4.5	35	10 US-09-776-491-14	Sequence 14, Appl
16	108	4.5	35	10 US-09-776-490-13	Sequence 13, Appl
17	108	4.5	35	10 US-09-776-490-15	Sequence 13, Appl
18	108	4.5	35	10 US-09-776-491-13	Sequence 13, Appl
19	108	4.5	35	10 US-09-776-491-15	Sequence 15, Appl

20	107.5	4.5	548	9 US-10-047-542-78	Sequence 78, Appl
21	106	4.4	35	10 US-09-776-490-9	Sequence 9, Appl1
22	106	4.4	35	10 US-09-776-491-9	Sequence 9, Appl1
23	105.5	4.4	951	9 US-09-924-097-15	Sequence 15, Appl
24	104	4.3	35	10 US-09-776-490-18	Sequence 18, Appl
25	104	4.3	35	10 US-09-776-491-18	Sequence 18, Appl
26	103	4.3	35	10 US-09-776-490-11	Sequence 11, Appl
27	103	4.3	35	10 US-09-776-491-11	Sequence 11, Appl
28	102	4.2	35	10 US-09-776-490-8	Sequence 8, Appl1
29	102	4.2	35	10 US-09-776-490-10	Sequence 10, Appl
30	102	4.2	35	10 US-09-776-491-8	Sequence 8, Appl1
31	102	4.2	35	10 US-09-776-491-10	Sequence 10, Appl
32	101.5	4.2	376	9 US-09-738-626-6115	Sequence 6115, Ap
33	101	4.2	17	10 US-09-776-490-44	Sequence 44, Appl
34	101	4.2	17	10 US-09-776-490-45	Sequence 44, Appl
35	101	4.2	17	10 US-09-776-491-44	Sequence 44, Appl
36	101	4.2	17	10 US-09-776-491-45	Sequence 45, Appl
37	100	4.2	35	10 US-09-776-490-17	Sequence 17, Appl
38	100	4.2	35	10 US-09-776-491-17	Sequence 17, Appl
39	97	4.0	3069	9 US-09-712-363-246	Sequence 246, App
40	95	3.9	17	10 US-09-776-490-46	Sequence 46, Appl
41	95	3.9	17	10 US-09-776-491-46	Sequence 46, Appl
42	94.5	3.9	282	10 US-09-925-301-1362	Sequence 1362, Ap
43	93	3.9	35	10 US-09-776-490-16	Sequence 16, Appl
44	93	3.9	35	10 US-09-776-491-16	Sequence 16, Appl
45	91	3.8	917	10 US-09-815-242-5603	Sequence 5603, Ap

#### ALIGNMENTS

RESULT 1  
US-09-815-242-10253  
; Sequence 10253, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA, 011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10253  
; LENGTH: 453  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10253  
Query Match 30.5%; Score 735; DB 10; Length 453;  
Best Local Similarity 36.4%; Pred. No. 1.8e-57;

Matches	165:	Conservative	64:	Mismatches	174:	Indels	50:	Gaps	8:
Qy	20	IGHDELEFQHELEKTTIFARMMLFLTHDSLIPARGDYTAAMGIDEVIVSRONDGSIRAFLN	79						
Db	24	ITDIPDIYOLETERIEGRICWMLFANESQIPKQDGFNTYMGEDAVVVRKQDSIKAFLN	83						
Qy	80	VRNRHGKTVSYEAGNAKCFVCSYHGMWCGSNGELOSYPEKDLVGESLNNKCKLKEVA	139						
Db	84	QCRHRMARVSYADCCGTRAFCTCYHMSYXINGELIDVLEPPRAVPQGLCKSHMGINLNP	143						
Qy	140	RSEFSEHGFYGEFDDPARPLMDYLGAANYLEPMF-KHSGLELVGPRPKVVIKANWKP	198						
Db	144	CVESTKGLIFGWMDSAPBLRDYLDGIANTYLDGMIDRREGCTFIYGVQKWVINCWKEP	203						
Qy	199	AENFVGDAYHVGWTHASSL-----RSGESIFSSLAGNALP-----	234						
Db	204	ABQFASDQYHALFHSASAVOLGAKDDGSDKRLDG-----GQTRAPWETAKDALQFG	256						
Qy	235	--PEGAGLQMTSKYSGMGVLYMDGYSGVASADLYPPLMAFGAKOERLKEIGDVARILY	292						
Db	237	QDGHSGEFEETKEKPRANVWV--DGAVSSYRRETYAE-----ABQRLGEVALRL	303						
Qy	293	RSHLQCTVPPNNSMLTCSGVFWKPNIDANTFEEVWYIAVEKDMPEDEKRLRLADSVQPT	352						
Db	304	AGHNN--ITPTLSWLNGTITLTKVWHRGDOYEWNACFTTDRAASDEVAAENSTRAF	361						
Qy	353	GPAGFESDNDNMETASONGRKYOSRDSLSNLGFGEDVYGDAYVPGVSKSAIGETS	412						
Db	362	GPAGFEQDDSENWCEIQKLTGKHRRARNSKCLCMLGEGEKRRDQIPGIT--NYIFSETA	420						
Qy	413	YRGFYRAVOAHYSSSNWAEFEHASSWHTLTK	445						
Db	421	ARGMYQRMADLLSSSEWQVLELDKTAAYAQOEVNK	453						
RESULT 2									
US-09-815-242-11692									
Sequence 11692, Application US/09815242									
Patent No. US2002061569A1									
GENERAL INFORMATION:									
APPLICANT: Haselbeck, Robert									
APPLICANT: Ohlsen, Kari L.									
APPLICANT: Zyskind, Judith W.									
APPLICANT: Wall, Daniel									
APPLICANT: Trawick, John D.									
APPLICANT: Carr, Grant J.									
APPLICANT: Yamamoto, Robert T.									
TITLE OF INVENTION: Identification of Essential Genes in									
FILE REFERENCE: ELITRA.011A									
CURRENT APPLICATION NUMBER: US-09/815,242									
CURRENT FILING DATE: 2001-03-21									
PRIOR APPLICATION NUMBER: 60/191,078									
PRIOR FILING DATE: 2000-03-21									
PRIOR APPLICATION NUMBER: 60/206,848									
PRIOR FILING DATE: 2000-05-23									
PRIOR APPLICATION NUMBER: 60/207,727									
PRIOR FILING DATE: 2000-05-26									
PRIOR APPLICATION NUMBER: 60/242,578									
PRIOR FILING DATE: 2000-10-23									
PRIOR APPLICATION NUMBER: 60/253,625									

[illegible]

```

RESULT 3
US-09-738-626-6140
; Sequence 6140, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HAROHICO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patentin ver. 3.0
; SEQ ID NO 6140
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6140

Query Match          15.6%; Score 375; DB 9; Length 490;
Best Local Similarity 26.3%; Pred. No. 1.8e-25;
Matches 118; Conservative 66; Mismatches 215; Indels 50; Gaps 15;

OY      20  IHGDELFQHELTIFARRWLFETHDSLIPAGDYVYAKMGIDEVYISRONDSIAFLN 79
      |  | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      27  IFTDPELFTELEMHRIEGWVYLAHESQIPNAGDYFTTYIGROPIMLTRKSDDTLNLIN 86
OY      80  VCRHGGTIVSVAGNAKGFVCSYGMGFGSNGELQSVPEKQ-LYGESLNKK-CIGLKE 137

```



Db 87 ACSHRGAMLRRTDNRTTLTLPFHGWTFSNDGALLKVKKEKAGAYPENRTDGSIDLRR 146  
QY 138 VARHESHGFIYCGFDEAPRLMDYLGDAAMYLEPMFKHS-GGLELVPPGKXVIRANMK 136  
Db 147 VPFRESTRGLFSLNDNDVSLSEHLGDRTVLMDLDOSPEGLLEVLRGSSTYYDGMNK 206  
QY 197 APAENFVGADYHVGWTH-----ASSLRSGESIFSSLAGNAALPPGAGLQMTSKYSGSM 250  
Db 207 LQENH-GADYHVSSTHWNATTSRGTGSEANETKAMAGWKGOGGSEYPIGCHML 255  
QY 251 GVLMDGYSVHSDVLPBELMAFGAKOERLNKEIGDVAR-IRSHLNTCTVFPNNSML-T 308  
Db 266 --LMMWNGNEDRPLF-----ERRDEKKEFEKGEKGFMYGASHNLCIYPNVYLMQ 315  
QY 309 CSQVFKWNPIDANTTEWYTAIVEKMDPEDLKRRLADSVQRTGPRGFESDNDNMET 368  
Db 316 FSSQIRIRIPISVDQTEVTLTYCIAPEGSAEATANRIROYEDFNATGMATPDLEEFRS 375  
QY 369 ASONGKKYQSRD--SDLSNLGFGEDYVDAYVPGVYCK-----SAIGETSYRGFY 417  
Db 376 CQ---KTYQASAPFWMND--TRGLGHQYCG---PNEVAKGLGMNVLSSGARTDEDEGLY 426  
QY 418 -----RAYQAHVSSSNMAEFELHASST 438  
Db 427 PIGHGFWEHLMQEAUVNKOSIKEKELADDT 455

## RESULT 4

US-09-815-242-5097  
; Sequence 5097, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011a  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5097  
; LENGTH: 424  
; TYPE: prt  
; ORGANISM: Pseudomonas aeruginosa  
US-09-815-242-5097

## Query Match

Best local Similarity 14.0%; Score 336.5; DB 10; Length 424;  
Matches 114; Conservative 70; Mismatches 178; Indels 55; Gaps 16;

QY 25 ELFOHEIKTIFARNWLELTHDSLIPAPGDYVYAKMGIDEIVSRONDGSIRAFVNCRRH 84

Db 24 ELARRELEHIFDSDSWLYAAHLSLRPEDEITRDVGRLIIQRADGEPAYVLNCAHR 83  
QY 85 GKTLVSEAGNAGPFCSTYHNGFGSNGELQSVFPEEDLYGSELNKKC---LGLKEVARY 141  
Db 84 GARVCARQNSORFPCPHGWTHSHGLSLGP-DKAAQHA--GQCHBELSLTQKHA 140  
QY 142 ESFHEGLYCGFDEAPRLMDYLGDAAMYLEPMFKHS-GGLELVPPGKXVIRANMKAPAE 200  
Db 141 -VYRNLFTHYGRQASLETLTGQAKDYIDLICDQSEALELTIIPGFHSHIKNMILLAE 199  
QY 201 NFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPPGAGLQMTSKYSGSMVLMDYSGV 260  
Db 200 NGV-DAYHLPFAKRYLEYLNTL-----CTDESHKRHRGRG-EALGNHALI 244  
QY 261 HS-----ADLVEP-LMAFGAKOERLNKEIGDVAR-IRSHLNTCTVFPN--- 303  
Db 245 ISGPPSTGRPIAYWSPLPEALKPSIAAKFERLVERGGARADIAHTNKSILFIFNLVI 304  
QY 304 NSMLTSGVFKWNPIDANTTEWYTAIVEKMDPEDLKRRLADSVQRTGPRGFESDND 363  
Db 305 NDILGLN--IRSEFPYRADSVTVWAGAFADETREBARARINGLISFTGPGGFTGPPDY 362  
QY 364 DNNETASQNGKKYQSRDSDLSNLGFGEDYVDAYVPGVYKSAIGETSYRGFYRAY 420  
Db 363 EILESQ---RAYAH-----AALGYSDFSRG---MGPATRRHVDQENRGCFWREM 406

## RESULT 5

US-09-776-490-19  
; Sequence 19, Application US/09776490  
; Patent No. US20010012886A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; Gray, John  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009  
; CITY: Charlotte  
; STATE: No. 280010012886A1th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/776,490  
; FILING DATE: 02-Feb-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/810,009  
; FILING DATE: 04-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spruill, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: 5718-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; TELEX: 575102  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-490-19

Query Match  
Best Local Similarity 7.7%; Score 185; DB 10; Length 35;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGNGFGSNGE 113  
|||||  
Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGNGFGSNGK 35

RESULT 6  
US-09-776-491-19  
Sequence 19, Application US/09776491  
Patent No. US20010013135A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776.491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810.009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102

INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-491-19

Query Match  
Best Local Similarity 7.7%; Score 185; DB 10; Length 35;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGNGFGSNGE 113  
|||||  
Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGNGFGSNGK 35

RESULT 7  
US-09-776-490-20  
Sequence 20, Application US/09776490  
Patent No. US20010012886A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776.490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810.009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-490-20

Query Match  
Best Local Similarity 7.2%; Score 174; DB 10; Length 35;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCRHRGKTLVSEAGNAKGFVCSYHGNGFGSNGE 113  
|||||  
Db 1 NVCRHRGKTLVSEAGNAKGFVCSYHGNGFGSNGK 35

RESULT 8  
US-09-776-491-20  
Sequence 20, Application US/09776491  
Patent No. US20010013135A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS

NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/776,491  
 FILING DATE: 02-Feb-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/810,009  
 FILING DATE: 04-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Spruill, W. Murray  
 REGISTRATION NUMBER: 32,943  
 REFERENCE/DOCKET NUMBER: 5718-4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-881-3140  
 TELEFAX: 919-881-3175  
 TELEX: 575102  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
 US-09-776-491-20

Query Match 7.2%; Score 174; DB 10; Length 35;  
 Best Local Similarity 85.7%; Pred. No. 2.9e-09;  
 Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 79 NVCRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGE 113  
 DB 1 NVCRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGK 35

RESULT 9  
 US-09-776-490-21  
 Sequence 21, Application US/09776490  
 Patent No. US20010012886A1  
 GENERAL INFORMATION:  
 APPLICANT: Briggs, Steven P.  
 Gray, John  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
 CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
 STREET: P.O. Drawer 34009  
 CITY: Charlotte  
 STATE: NO. US20010012886A1th Carolina  
 COUNTRY: USA  
 ZIP: 28234  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/776,490  
 FILING DATE: 02-Feb-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/810,009  
 FILING DATE: 04-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Spruill, W. Murray  
 REGISTRATION NUMBER: 32,943  
 REFERENCE/DOCKET NUMBER: 5718-4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175  
 TELEX: 575102  
 INFORMATION FOR SEQ ID NO: 21:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
 US-09-776-490-21

Query Match 7.0%; Score 168; DB 10; Length 35;  
 Best Local Similarity 80.0%; Pred. No. 9.8e-09;  
 Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 79 NVCRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGE 113  
 DB 1 NVCRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGK 35

RESULT 10  
 US-09-776-491-21  
 Sequence 21, Application US/09776491  
 Patent No. US20010013135A1  
 GENERAL INFORMATION:  
 APPLICANT: Briggs, Steven P.  
 Gray, John  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
 CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
 STREET: P.O. Drawer 34009  
 CITY: Charlotte  
 STATE: NO. US20010013135A1th Carolina  
 COUNTRY: USA  
 ZIP: 28234  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/776,491  
 FILING DATE: 02-Feb-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/810,009  
 FILING DATE: 04-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Spruill, W. Murray  
 REGISTRATION NUMBER: 32,943  
 REFERENCE/DOCKET NUMBER: 5718-4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-881-3140  
 TELEFAX: 919-881-3175  
 TELEX: 575102  
 INFORMATION FOR SEQ ID NO: 21:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
 US-09-776-491-21

Query Match 7.0%; Score 168; DB 10; Length 35;  
 Best Local Similarity 80.0%; Pred. No. 9.8e-09;  
 Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Oy 79 NVCRHRRGTLVSEAGNAKGFYCSYHGWFSGNGE 113  
|||||:-||| ||||| |||||:-|||  
Db 1 NVCRRHRGKTLVDAAEGNAKGPRYCGYHGWGYSGNGK 35

```

RESULT 11
US-08-976-063C-4
; Sequence 4, Application US/08976063C
; Publication No. US20020182697A1
GENERAL INFORMATION:
APPLICANT: Alexander Stelubuchel; Horst Priefert; Jürgen Rabenhorst
TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF
TITLE OF INVENTION: CONFERYL ALCOHOL, CONFERYLALDEHYDE, FERULIC ACID, VANILLIN A
TITLE OF INVENTION: ACID AND THEIR USE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SPRUNG KRAMER SCHAEFER & BRISCOE
STREET: 660 White Plains Road
City: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage
COMPUTER: HP VECTOR
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976,063C
FILING DATE: 21-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 196 49 655.1 (Germany)
FILING DATE: 29-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 9998-CAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 354 amino acids
Type: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-063C-4

```

```

Query Match      4.9%; Score 118.5; DB 1; Length 354;
Best Local Similarity 23.2%; Pred. No. 0.0067;
Matches 56; Conservative 40; Mismatches 78; Indels 67; Gaps 14.

QY 34 IFARN-WLFLTHDSLIPAGDYVTAKMGIDEVIVSRONDSIRAFILNVCRRHGKTLVSV 91
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MFPKNAMVACTPPDLIAKP--LGRJONEKIVFYRPGEGRAVAVEDCPHRAPL--- 54

QY 92 EAGNAKGV-----CSYHMGVFGSNGELQSVPEFKDLGESLNKKCLG--LKEARVE 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 ----SLGFVRDGKILCGHYHGLEMGCEGKTLAMP-----GQRVQGFPCIKSYA-VE 99

QY 143 SFHGFIY---GCFPDQAPPLLDYLGDAAMYLEPMFKHSGLELVGPPGKVITAKMKAAPA 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 ERYGFIVWPPDRLADPALIHHL---EMADNPENAVAGGL-----YHACDYRLMI 148

QY 200 ENFGDAHVAGVMTTHASSLR-----SGRSIFSS-LAGNAALP-----EGAG 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 DNLN-DLTHETHYVHASSIGQKEIDEAPVSTRYEGDVTISRMDVMNMAPPFVRAALRGNG 207

QY 240 L 240
   |
Db 208 L 208

```

RESULT 12  
 US-09-776-490-12  
 Sequence 12, Application US/09776490  
 Patent No. US20010012886A1  
 GENERAL INFORMATION:  
 APPLICANT: Briggs, Steven P.  
 Johnal, Gurmukh S.  
 Gray, John  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
 CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
 STREET: P.O. Drawer 34009  
 City: Charlotte  
 STATE: NO. US20010012886A1th Carolina  
 COUNTRY: USA  
 ZIP: 28234  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentln Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/776,490  
 FILING DATE: 02-Feb-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/810,009  
 FILING DATE: 04-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Spruill, W. Murray  
 REGISTRATION NUMBER: 32,943  
 REFERENCE/DOCKET NUMBER: 5718-4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-881-3140  
 TELEFAX: 919-881-3175  
 TELEX: 575102  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
 IS-09-776-490-12

```

Query Match: 4.6%; Score 110; DB 10; Length 35;
Best Local Similarity 52.9%; Pred. No. 0.0014;
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0.

Qy 79 NVCHRNGKTLVSEAGNAGKGFVCSYHGMGFGSG 112
| | | | | : | | | | | | | | | : |
Db 1 NQCRHRGMRICRADAGNAKAPTCSTHGNAYPDAG 34

RESULT 13
US-09-776-491-12
; Sequence 12, Application US/09776491
; Patent No. US20010013135A1
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; Johnal, Gurmukh S.
; Gray, John
;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009

```

CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-491-12  
Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0014;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;  
QY 79 NVCRRHGKTLVSEAGNAGKGVCSYHGMPGSGNG 112  
DB 1 NVCRRHGKTLVSEAGNAGKGVCSYHGMPGSGNG 34  
RESULT 14  
US-09-776-490-14  
Sequence 14, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-776-490-14  
Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0017;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
QY 79 NVCRRHGKTLVSEAGNAGKGVCSYHGMPGSGNG 112  
DB 1 NVCRRHGKTLVSEAGNAGKGVCSYHGMPGSGNG 34  
RESULT 15  
US-09-776-491-14  
Sequence 14, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-776-491-14

Query Match 4.5%; Score 109; DB 10; Length 35;

Best Local Similarity 52.9%; Pred. No. 0.0017;

Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHREKTLVSEAGNAKGFVCSYHGWFSGSNG 112  
| | | | | : | | | | | : | |  
Db 1 NQCRHRGMRICRADGNAKSPFTCSYHGWAYDSAG 34

Search completed: January 27, 2003, 09:20:56  
Job time : 8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:56:07 ; Search time 10.1667 Seconds  
(without alignments)  
1299.432 Million cell updates/sec

Title: US-09-843-250-34

Perfect score: 2409

Sequence: 1 MNYNKKILVSESGLSQKHLI.....AEFHASSTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2.6/ptodata/1/1aa/PTCUTS.COMB.pep:\*  
6: /cgn2.6/ptodata/1/1aa/Backfilltest.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	224	9.3	446	US-09-004-393B-4	Sequence 4, Appli
2	223	9.3	439	US-09-004-393B-2	Sequence 2, Appli
3	185	7.7	35	US-08-810-009-19	Sequence 19, Appli
4	174	7.2	35	US-08-810-009-20	Sequence 20, Appli
5	168	7.0	35	US-08-810-009-21	Sequence 21, Appli
6	113.5	4.7	379	US-09-028-934-36	Sequence 36, Appli
7	110	4.6	35	US-08-810-009-12	Sequence 12, Appli
8	109	4.5	35	US-08-810-009-14	Sequence 14, Appli
9	108	4.5	35	US-08-810-009-13	Sequence 13, Appli
10	108	4.5	35	US-08-810-009-15	Sequence 15, Appli
11	108	4.5	622	US-09-311-626B-4	Sequence 13, Appli
12	106.5	4.4	432	US-08-809-326A-16	Sequence 16, Appli
13	106.5	4.4	649	US-08-809-326A-15	Sequence 15, Appli
14	106	4.4	35	US-08-810-009-9	Sequence 9, Appli
15	104	4.3	35	US-08-810-009-18	Sequence 18, Appli
16	103	4.3	35	US-08-810-009-11	Sequence 11, Appli
17	102	4.2	35	US-08-810-009-8	Sequence 8, Appli
18	102	4.2	35	US-08-810-009-10	Sequence 10, Appli
19	101	4.2	17	US-08-810-009-44	Sequence 44, Appli
20	101	4.2	17	US-08-810-009-45	Sequence 45, Appli
21	100	4.2	35	US-08-810-009-17	Sequence 17, Appli
22	99.5	4.1	256	US-09-325-932A-57	Sequence 17, Appli
23	96	4.0	563	US-09-134-001C-4800	Sequence 46, Appli
24	95	3.9	17	US-08-810-009-46	Sequence 46, Appli
25	94.5	3.9	350	US-08-828-922-1	Sequence 1, Appli
26	94	3.9	560	US-08-814-052-6	Sequence 6, Appli
27	94	3.9	560	US-08-812-829-6	Sequence 6, Appli

28	94	3.9	5588	4	US-09-036-987A-6	Sequence 6, Appli
29	94	3.9	5588	4	US-09-370-700-6	Sequence 6, Appli
30	93	3.9	35	4	US-08-810-009-16	Sequence 16, Appli
31	93	3.9	421	1	US-09-239-303-2	Sequence 2, Appli
32	92.5	3.8	525	1	US-08-348-891A-2	Sequence 2, Appli
33	92.5	3.8	525	2	US-08-905-817-2	Sequence 2, Appli
34	92	3.8	1087	2	US-08-311-8	Sequence 8, Appli
35	92	3.8	1087	2	US-08-353-485-8	Sequence 8, Appli
36	92	3.8	1358	2	US-08-570-311-27	Sequence 27, Appli
37	91.5	3.8	1785	4	US-09-341-587-3	Sequence 3, Appli
38	91	3.8	616	2	US-08-545-745-2	Sequence 2, Appli
39	90.5	3.8	345	4	US-09-362-123A-2	Sequence 2, Appli
40	90.5	3.8	346	4	US-09-149-476-493	Sequence 4, Appli
41	90.5	3.8	774	4	US-09-276-400-7	Sequence 4, Appli
42	90.5	3.8	774	4	US-09-448-076-7	Sequence 7, Appli
43	90.5	3.8	774	4	US-09-702-572-7	Sequence 7, Appli
44	90.5	3.8	997	1	US-08-232-540-1	Sequence 1, Appli
45	90.5	3.8	997	1	US-08-428-949A-1	Sequence 1, Appli

#### ALIGNMENTS

```

RESULT 1
US-09-004-393B-4
; Sequence 4, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
; APPLICANT: Rathinasabapathi, Bala
; APPLICANT: Burnett, Michael
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; TITLE OF INVENTION: Plants Transformed Therewith
; FILE REFERENCE: US-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-004-393B-4

Query Match          9.3%; Score 224; DB 4; Length 446;
Best Local Similarity 27.4%; Pred. No. 3.9e-13;
Matches 58; Conservative 37; Mismatches 97; Indels 20; Gaps 6;

QY 11 ESGLSQKHLIHGDELLFQHELTFTFARMMLFTTHSLIPAPGDYVTAKMGIDEIVYSRON 70
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 99 EDALPPSTWTEPAPFTLSVLEERITFKGMQVAGYSEQYKKNQYFTGSLGNVEYVSRDG 158
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 71 DGSIRAFILNVCRRKRTFLSVAGNAKGFVCSYHSGFSGNGLQSVPEFKDLGESINK 130
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 159 OGELHAFINVCVHRA-SILAGSGKSKSCFVCPYHGWYGLDGLAKA--SKATEFQNDP 215
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 KCLGKEVARVESFTGFTYGCPCDQAPPLMD---YLDDA-----AWLEPFKHSGLG 181
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 KELGIAPL-KVAEMCPFLISLDRSLDANADVGTWICKSAEDVKAHAFDPLKFTHRSE 274
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 LVGPFGKVVIRANKKAPAEENFVGDAYHGWTH 213
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 F-----PMECNKVFPCDNLIDSSYHVPYAH 299
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-09-004-393B-2
; Sequence 2, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew

```





Sequence 21, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmuh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437ch Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-21

Query Match  
Best Local Similarity 7.0%; Score 168; DB 4; Length 35;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCRRHRTLVSYEAGNAGKGFVCSYHGMFGSGNGE 113  
DB 1 NVCRRHRTLVDAEAGNAGKGFVCGHGMFGSGNGK 35

RESULT 6  
US-09-028-934-36  
Sequence 36, Application US/09028934  
Patent No. 6117670.  
GENERAL INFORMATION:  
APPLICANT: Ligon, James M.  
APPLICANT: Hill, Dwight S.  
APPLICANT: Lam, Steven T.  
APPLICANT: Hammer, Philip E.  
APPLICANT: van Bee, Karl-Heinz  
APPLICANT: Kirner, Sabine  
APPLICANT: Young, Thomas R.  
TITLE OF INVENTION: Pyrolytic Biosynthesis Genes and Uses  
TITLE OF INVENTION: Pyrolytic  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6117670artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/028,934  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/729,214  
FILING DATE: 09-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 379 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-028-934-36

Query Match  
Best Local Similarity 4.7%; Score 113.5; DB 3; Length 379;  
Matches 95; Conservative 54; Mismatches 149; Indels 171; Gaps 26;

QY 20 IHGDELEPQ-----HKLKIPARNW-LFTHSDLPAPBDYTAAMGIDEVYSRONDS 73  
DB 5 IHQEPRIQASGVNDLTQTOTASWVAMRSDALRGKP--VAIKLFGQPLVAMRDGGGR 61

QY 74 IRAFLNVCRRHRTLVSYEAGNAGKGFV-----CSYHGMFGSGNGELQSVPEKDLGHS 127  
DB 62 PYVMERYCCHLASL-----AKGVYEGCICQCPHNRKYSTGACSHV-----GHS 108

QY 128 LNKCKICGLKEVARVESE-----HGYGCFDQAP---PLMDYLGDAWY 169  
DB 109 -----TEVRLPEIPPTARQSVYPVMERYGFVWVYGTAKPLPLPEMEAESESSEH 160

QY 170 LPMFHSGLGELVGPQGVYIKANKKAPAEPMVGDATVHGTHASLSRSGE-SIF---S 225  
DB 161 QSLRAVEYETTTSVL---RII-----ENFY-DAQHAAPVHOLPLISAFELKLFDESS 206

QY 226 SLAGNALPPE-----GAGLQW-TSKYSGMGVL-----WDGYGVHSDLV 266  
DB 207 PPGGEALARDQAMFGAGIDFHDVRFGLGVISRTLGISMARMOLHPFGYIG----- 259

QY 267 PELMAFGAKQERLNKEIGDVARIRYRSHLNCTVPPNNSMLTCSGVFKWYNDIDANTTEV 326  
DB 260 -----GCIMTVSLD---GDVKYRL---LQCV-----TEVDKEETVM 289

QY 327 WTYATVEKMPEDLKRRLADSV---QRTGPRAGF---WESDNDNMETASONGKKYQSR 379  
DB 290 HMLAIRKKG--DGVVRSANFTLYGLQTTAAAGYVAIVNSKAD-----GGGAFSKY 340

QY 380 DSDLSNLGFGEDVYGDVAYPGVVKSAIGTSYSGFYRAYOAHSSN 428  
DB 341 DDLIL-----KYRAFIRRWYKVALEN 362

RESULT 7  
US-08-810-009-12  
Sequence 12, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.

APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-12

Query Match 4.6%; Score 110; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.00071;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRHGRKTVSEAGNAKGFVCSYHGWFSGSNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NQCRHGRICRADGNAKAPTCSTHGWAYDTAG 34

RESULT 8  
US-08-810-009-14  
Sequence 14, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-14

Query Match 4.5%; Score 109; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.00088;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHGRKTVSEAGNAKGFVCSYHGWFSGSNG 112  
| | | | | : | | | | | | | | | : | |  
Db 1 NQCRHGRICRADGNAKAPTCSTHGWAYDSAG 34

RESULT 9  
US-08-810-009-13  
Sequence 13, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-13

Query Match 4.5%; Score 108; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0011;

CURRENT FILING DATE: 1999-05-13  
 PRIOR APPLICATION NUMBER: 0608/98  
 PRIOR FILING DATE: 1998-05-01  
 PRIOR APPLICATION NUMBER: 60/084,358  
 PRIOR FILING DATE: 1998-05-05  
 NUMBER OF SEQ ID NOS: 61  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 4  
 LENGTH: 622  
 TYPE: PRF  
 ORGANISM: Bacillus licheniformis  
 US-09-311-626B-4

Query Match            4.5%; Score 108; DB 4; Length 622;  
 Best Local Similarity 21.6%; Pred. No. 0.09; Mismatches 137; Indels 138; Gaps 23;  
 Matches 90; Conservative 52;

```

QY      84 RGLTVSYEAG--NAKGFVCS-----YHGWFSGNGELASVPPEK-----DLYG 125
       :|: |: | ||: |: | | | | | | | | | | | | | | | | | | | | | |
Db     245 KGVGNANADRYMNOGRHLSCEPYLTFVKG--DTGAELFTVTNEPARGNVADMGDGYG 301
QY      126 ESLNKKCGLG-----KEVARVSFFHGFIYG-----CFDQEAPLMD 161
       :|: |: | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     302 NRVDREFLAGVALDGERPSFYWARGCYTRTYLVAVNFPG---GKLTKLMTFDSDAPNGA 358
QY      162 YLDGAAMYLEPMFKHSGGLELVGPFGKVVIKANMKAPAEENFVGDAYHYGMFHASSLRSGE 221
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     359 YAGQS-----NHSLSVADVDDGGKDEIITYIGAAVVDHDCKG-LYSTGWGHGDMATHGN 409
QY      222 SIFSSLAGNALPPRGAGLOMTSKYSGMGVLMIDGYSCVHSADLVPELMAFQAKOERL- 280
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     410 -----LDPSRPGLVFQ-----VHENSNSPYGLSPRDAATGKKII 443
QY      281 -----NKEIG-DVRAIRYSHLCTFPNNSLTDCSVEFKVWPNDATT-EVMYTAIVE 333
       | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     444 WGVHAGKRVGGMADDIPRYEGAEVWANGSLYTKAGY-KIGNILPSTNTGIM----- 496
QY      334 KDMPEDLKRLADSVORTTGPAFWESDDND--NMETAS----ONG-KKYOSRDSLLSN 386
       ||: ||: ||: | | | | | | | | | | | | | | | | | | | | | |
Db     497 -WDPDLOGRELIDS---NRIDKWIDYONSRTVINLTFSASANNGTKATPSLOADILGD 549
QY      387 -----LGRGED-----YYGDAVYRGVYGSALG--ESTYRGFY 417
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     550 WRREYVWRABESSSELRITTTDYTEHRMYTLMDHAVYRLGIAMONVGYNOPPHGTGY 606

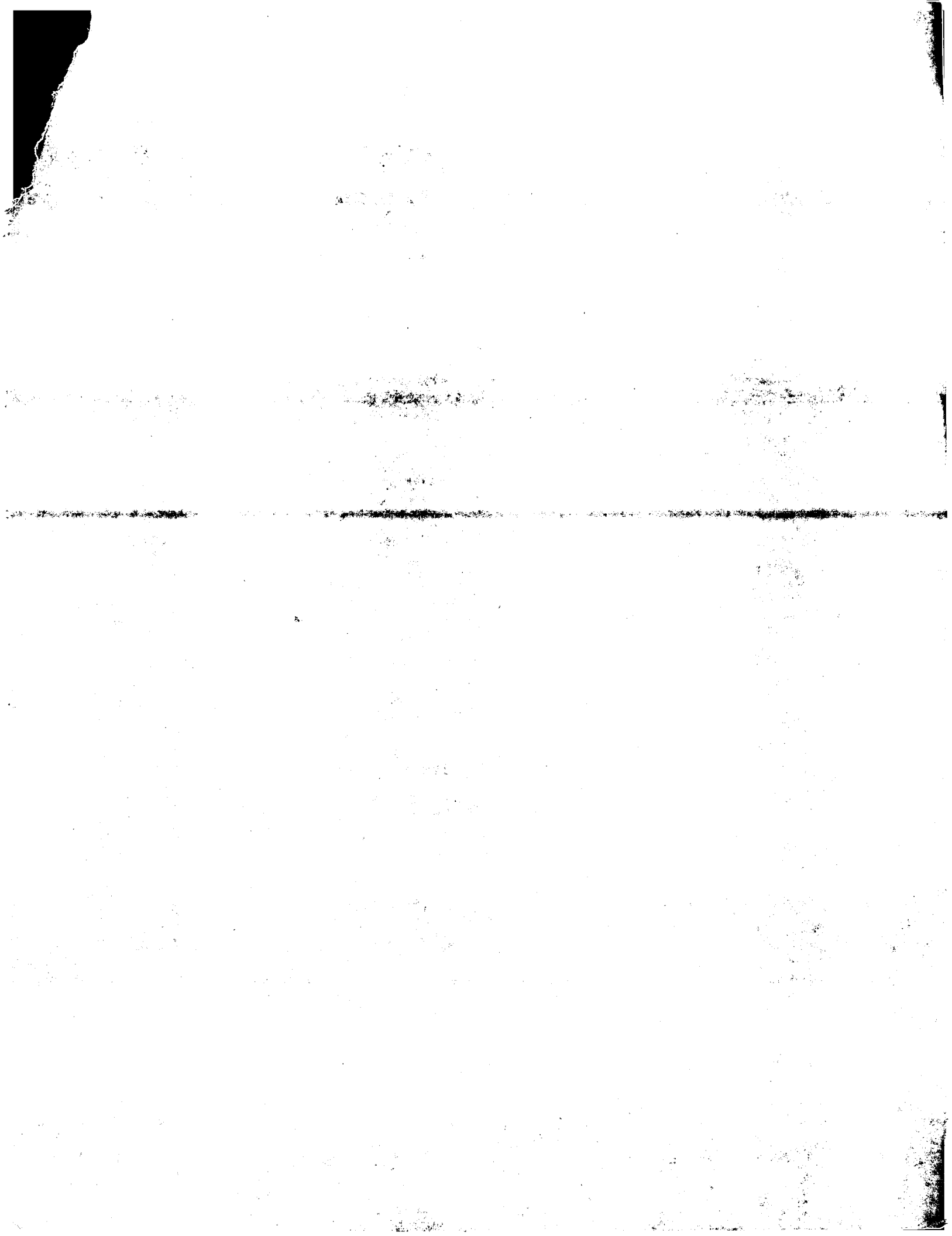
```

RESULT 12  
 US-08-809-326A-16  
 Sequence 16 Application US/08809326A  
 Patent No. 6165478  
 GENERAL INFORMATION:  
 APPLICANT: Izutsu, Hiroshi  
 APPLICANT: Obara, Kazuhiko  
 APPLICANT: Matsumoto, Akira  
 TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
 TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
 TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMA  
 TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
 TITLE OF INVENTION: ANTI-BODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMEN  
 TITLE OF INVENTION: OF ANTI-BODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
 TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
 TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
 TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
 TITLE OF INVENTION: PNEUMONIAE GENE  
 NUMBER OF SEQUENCES: 31  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pennie & Edmonds LLP  
 STREET: 1155 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036-2711  
 COMPUTER READABLE FORM:



; Patent No. 6211437  
; GENERAL INFORMATION:

Search completed: January 27, 2003, 09:04:55  
Job time : 11.1667 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:54:32 ; Search time 25.8333 Seconds  
(without alignments)  
3581.232 Million cell updates/sec

Title: US-09-843-250-34  
Perfect score: 2409  
Sequence: 1 MNYNKKIIVSESGLSQKHLI.....AEFEHASSWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp-archaea:\*
- 2: sp-bacteria:\*
- 3: sp-fungi:\*
- 4: sp-human:\*
- 5: sp-invertebrate:\*
- 6: sp-mammal:\*
- 7: sp-mhc:\*
- 8: sp-organelle:\*
- 9: sp-phage:\*
- 10: sp-plant:\*
- 11: sp-rodent:\*
- 12: sp-virus:\*
- 13: sp-vertebrate:\*
- 14: sp-unclassified:\*
- 15: sp-virus:\*
- 16: sp-bacteriap:\*
- 17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2330	96.7	449	2	Q938R9 pseudomonas
2	2280	95.1	449	2	Q52142 pseudomonas
3	2206	91.6	449	2	Q92173 pseudomonas
4	2191	91.0	447	2	Q915D2 pseudomonas
5	2187	90.8	447	2	Q52382 pseudomonas
6	2071	86.0	447	2	Q8YUD4 pseudomonas
7	2051	85.1	447	2	P95564 pseudomonas
8	2022	83.9	447	2	Q8RT14 pseudomonas
9	1973.5	81.9	451	2	Q45695 pseudomonas
10	1490	61.9	277	2	Q9ETK2 pseudomonas
11	1488	61.8	277	2	Q9F5S4 pseudomonas
12	1486	61.7	277	2	Q9F5T3 pseudomonas
13	1484	61.6	277	2	Q9F5S8 pseudomonas
14	1483	61.6	277	2	Q9F5S9 pseudomonas
15	1451	60.2	277	2	Q9F5S3 pseudomonas
16	1405	58.3	277	2	Q9EUC7 pseudomonas

17	1403	58.2	277	2	Q9F5T8 pseudomonas
18	1402	58.2	277	2	Q9F5T5 pseudomonas
19	1399	58.1	277	2	Q9ETV3 pseudomonas
20	1399	58.1	277	2	Q9F5S5 pseudomonas
21	1396	57.9	277	2	Q9F5T6 pseudomonas
22	1396	57.9	277	2	Q9F5S7 pseudomonas
23	1393	57.8	277	2	Q9F5T4 pseudomonas
24	1393	57.8	277	2	Q9F5T2 pseudomonas
25	1393	57.8	277	2	Q9F5T1 pseudomonas
26	1393	57.8	277	2	Q9F5T0 pseudomonas
27	1393	57.8	277	2	Q9F5S6 pseudomonas
28	1390	57.7	277	2	Q9F5T7 pseudomonas
29	1384	57.5	277	2	Q9F5S2 pseudomonas
30	1353.5	56.2	450	2	Q92HH3 pseudomonas
31	1139	47.3	226	2	Q8YUM6 pseudomonas
32	1134	47.1	214	2	Q8VR23 pseudomonas
33	1134	47.1	214	2	Q8VL21 pseudomonas
34	1127	46.8	214	2	Q8VR24 pseudomonas
35	1123	46.6	214	2	Q8VR22 pseudomonas
36	111.5	46.1	455	2	Q93NA8 pseudomonas
37	1067	44.3	214	2	Q8VR25 pseudomonas
38	1043.5	43.3	437	2	Q9XG68 pseudomonas
39	1040	43.2	208	2	Q9F6B6 pseudomonas
40	1039	43.1	206	2	Q8YUM5 pseudomonas
41	1038	43.1	214	2	Q8VR21 pseudomonas
42	1038	43.1	303	2	Q8YUM2 pseudomonas
43	1016	42.2	297	2	Q93MA0 pseudomonas
44	1009	41.9	297	2	Q93M39 pseudomonas
45	1002	41.6	297	2	Q93M41 pseudomonas

## ALIGNMENTS

RESULT 1	Q938R9	PRELIMINARY:	PRT:	449 AA.
AC	Q938R9;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Naphthalene dioxygenase.			
OS	NAHAC.			
OS	Pseudomonas fluorescens.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=294;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Min K.-H., Ji S.-H.;			
RL	Submitted (OC-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AY048759; AAL07262.1; -.			
DR	InterPro; IPR001281; Rleske.			
DR	InterPro; IPR001663; Ring_hydroxyl_A.			
DR	Pfam; PF00355; Rleske; 1.			
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.			
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.			
KW	Dioxygenase.			
SQ	SEQUENCE 449 AA: 49608 MW; 5EFDDBE0282FE812 CRC64;			
Query Match	96.7%: Score 2330; DB 2: Length 449;			
Best Local Similarity	96.0%: Pred. No. 1.4e-170;			
Matches 431; Conservative 12; Mismatches 6; Indels 0; Gaps 0;				
QY	1 MNYNKKIIVSESGLSQKHLIHGDEELFQHELTIFARNMLFTTHDSLTPAPGDYATKMG 60			
DB	1 MNYNKKIIVSESGLSQKHLIHGDEELFQHELTIFARNMLFTTHDSLTPSPEDYATKMG 60			
QY	61 IDEVIVSRNDGSIATFLVNCRHGRKTLVSVAGNAKGVCSYHGMGSGNCELOSVPPE 120			
DB	61 IDEVIVSRNDGSIATFLVNCRHGRKTLVNAEAGNAKGVCSYHGMGSGNCELOSVPPE 120			
QY	121 KLYGSLNKKKCLGLKEVARVSPHGFYICGFDGEAPPLMDYLGDAWVLEPMFKHSGSL 180			

```

Db 121 KELGESLTKKCLGKEVARVESFHGFIYCGFDCEAPSLMDYLGAAMYLEPIFKHSGL 180
Qy 181 ELVPPGKVVIVKAMKKAENFVGDAHVGMTTHASSLSRGSIFSSLAGNALPREGAGL 240
Db 181 ELVPPGKVVIVKAMKKAENFVGDAHVGMTTHASSLSRGSIFSSLAGNALPREGAGL 240
Qy 241 QMTSKYSGMGLVDGYSGVHSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNCVY 300
Db 241 QMTSKYSGMGLVDGYSGVHSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNCVY 300
Qy 301 FPNNSMILTCGCVFVKWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVORTTGPAGWES 360
Db 301 FPNNSVLTCGCVFVKWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVORTTGPAGWES 360
Qy 361 DDNDMMETASONGKRYGSRDSDLSNLGFGEDVYDAYVPGVYKSAIGETSYRGFYRAY 420
Db 361 DDNDMMETASONGKRYGSRDSDLSNLGFGEDVYDAYVPGVYKSAIGETSYRGFYRAY 420
Qy 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449
Db 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449

```

## RESULT 2

PRELIMINARY: PRT: 449 AA.

```

ID Q52142 PRELIMINARY: PRT: 449 AA.
AC Q52142: 008194:
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Iron-sulfur protein large subunit.
GN PAHAC.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA Takizawa N.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RC MEDLINE=94209249; Pubmed=8157615;
RA TAKIZAWA N., Kaida N., Torigoe S., Moritani T., Sawada T., Satoh S.,
RA Kiyohara H.;
RT "Identification and characterization of genes encoding polycyclic
RT aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon
RT dihydrodiol dehydrogenase in Pseudomonas putida OUS82."
RL J. Bacteriol. 176:2444-2449(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA Noboru T., Toshiya I., Takashi S., Kazuhiro Y., Masamichi K.,
RA Yue-Wu W., Masao F., Hozoh K.;
RT "The molecular analysis of NAH7-type cluster located on the
RT chromosomes of Pseudomonas aeruginosa Pak1 and Pseudomonas putida
RT OUS82."
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: AB004059; BAA20391.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00848; Rieseke.1.
DR Pfam: PF00848; Ring_hydroxyl_A.1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA.1.
SO SEQUENCE 449 AA; 49361 MW; 80020F54AAB11EBA CRC64;

```

Query Match 95.1%; Score 2290; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 1.7e-167;

```

Matches 424: Conservative 13; Mismatches 12; Indels 0; Gaps 0;
Qy 1 MNYNKILVSESGLSQKHLHGDEELFOHELKTIETARNMLFLTHDSLIPAGDVTAKMG 60
Db 1 MNYNKILVSESGLSQKHLHGDEELFOHELKTIETARNMLFLTHDSLIPAGDVTAKMG 60
Qy 61 IDEVIVSRQNDGSTRAFNLVNCRHRRKTLVSEAGNAKGFVCSYHGWGFGSNGELQSVPE 120
Db 61 IDEVIVSRQNDGSTRAFNLVNCRHRRKTLVSEAGNAKGFVCSYHGWGFGSNGELQSVPE 120
Qy 121 KDLYESLNKCKCLGKEVARVESFHGFIYCGFDCEAPSLMDYLGAAMYLEPIFKHSGL 180
Db 121 KDLYESLNKCKCLGKEVARVESFHGFIYCGFDCEAPSLMDYLGAAMYLEPIFKHSGL 180
Qy 181 ELVPPGKVVIVKAMKKAENFVGDAHVGMTTHASSLSRGSIFSSLAGNALPREGAGL 240
Db 181 ELVPPGKVVIVKAMKKAENFVGDAHVGMTTHASSLSRGSIFSSLAGNALPREGAGL 240
Qy 241 QMTSKYSGMGLVDGYSGVHSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNCVY 300
Db 241 QMTSKYSGMGLVDGYSGVHSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNCVY 300
Qy 301 FPNNSMILTCGCVFVKWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVORTTGPAGWES 360
Db 301 FPNNSVLTCGCVFVKWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVORTTGPAGWES 360
Qy 361 DDNDMMETASONGKRYGSRDSDLSNLGFGEDVYDAYVPGVYKSAIGETSYRGFYRAY 420
Db 361 DDNDMMETASONGKRYGSRDSDLSNLGFGEDVYDAYVPGVYKSAIGETSYRGFYRAY 420
Qy 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449
Db 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449

```

## RESULT 3

PRELIMINARY: PRT: 449 AA.

```

ID Q92I73 PRELIMINARY: PRT: 449 AA.
AC Q92I73:
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Naphthalene dioxygenase Fe-S large subunit.
GN NAHAC.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AN10;
RC MEDLINE=99365311; Pubmed=10433976;
RA Bosch R., Garcia-Valdes E., Moore E.R.B.;
RT "Genetic characterization and evolutionary implications of a
RT chromosomally encoded naphthalene-degradation upper pathway from
RT Pseudomonas stutzeri AN10."
RL Gene 236:149-157(1999).
DR EMBL: AF039533; AAD02136.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00848; Rieseke.1.
DR Pfam: PF00848; Ring_hydroxyl_A.1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA.1.
KW Dioxygenase.
SO SEQUENCE 449 AA; 49802 MW; BA510FD1B6F66E63 CRC64;

```

Query Match 91.6%; Score 2206; DB 2; Length 449;  
Best Local Similarity 89.5%; Pred. No. 4.7e-161;  
Matches 402; Conservative 25; Mismatches 22; Indels 0; Gaps 0;



```

Db      1 MNKNKTLVTESGLTQKHLIHGDEELFORELETFIFARNMLFLTHDSLIPSGDYATAKKG 60
QY      61 IDEVIYSRONDGSIRAFNLVNCRRHCKTIVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120
Db      61 VDEVIVSRONDGSIRAFNLVNCRRHCKTIVLVAEKGNAKGFVCSYHGMFGSANGELQSVPE 120
QY      121 KDLVGSLLMKKCLGKLEVARVESFHGFIYGCPEQAPRLMDYLGDAWYLEPFFKHSGL 180
Db      121 KELYGALDKKCKGKLEVARVESFHGFIYGCPEQAPRLMDYLGDAWYLEPFFKHSGL 180
QY      181 ELVGPCKVIVIKANKKAPAEVFGDAYHVGMTHASSLSRSGESIFSSLAGNAALPREGAGL 240
Db      181 ELVGPCKVIVIKANKKAPAEVFGDAYHVGMTHASSLSRSGESIFSSLAGNAALPREGAGL 240
QY      241 QMTSKYSGMGVLMDSYGVSHADYLPBELMAFGAKOEKRLNEIGVRRARIYRSHLNCYV 300
Db      241 QMTSKYSGMGVLMDSYGVSHADYLPBELMAFGAKOEKRLNEIGVRRARIYRSHLNCYV 300
QY      301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRRLADSVORTTGPAGFWS 360
Db      301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRRLADSVORTTGPAGFWS 360
QY      361 DDNDNMETASONGKATYQSDSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
Db      361 DDNDNMETASONGKATYQSDSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
QY      421 QAHVSSNMAEFEBHASTWHTTELTKTTDR 449
Db      421 QAHVSSNMAEFEBHASTWHTTELTKTTDR 449

```

```

RESULT 4
Q9L5D2      PRELIMINARY:      PRT:      447 AA.
AC      Q9L5D2:      01-OCT-2000 (TREMBlrel. 15, Created)
DT      01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT      01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE      Polyaromatic hydrocarbon dioxygenase large subunit.
GN      PAHAC.
OS      Comamonas testosteroni (Pseudomonas testosteroni).
OC      Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX      NCBI_TaxID=285;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H;
RX      MEDLINE=21306905; PubMed=11414329;
RA      Moser R., Stahl U.;
RT      "Insights into the genetic diversity of initial dioxygenases from PAH-
      degrading bacteria."
RL      Appl. Microbiol. Biotechnol. 55:609-618(2001).
DR      EMBL; AF252550; AAF72976.1; -.
DR      HSSP; P23094; INDO.
DR      InterPro; IPR001281; Rleske.
DR      InterPro; IPR001663; Ring_hydroxyl_A.
DR      Pfam; PF00355; Rleske; 1.
DR      Pfam; PF00848; Ring_hydroxyl_A; 1.
DR      PRINTS; PR00090; RINGDIOXGNASE.
DR      PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW      Dioxygenase.
SQ      SEQUENCE 447 AA; 49548 MW; 2950EF36123A6F21 CRC64;

```

```

Query Match      91.0%; Score 2191; DB 2; Length 447;
Best Local Similarity 88.9%; Pred. No. 6,7e-160;
Matches 399; Conservative 27; Mismatches 21; Indels 2; Gaps 1;
QY      1 MNYNKILVSESGLSQKHLIHGDEELFQHELKTIIFARNMLFLTHDSLIPAPGDYATAKKG 60
Db      1 MIYEN--LVSEAGLQKHLIHGDKELFQHELKTIIFARNMLFLTHDSLIPSPDYATAKKG 58
QY      61 IDEVIYSRONDGSIRAFNLVNCRRHCKTIVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120

```

```

Db      59 VDEVIVSRONDGSVRAFLVNCRRHCKTIVLVAEAGNAKGFVCSYHGMFGSNGELQSVPE 118
QY      121 KDLVGSLLMKKCLGKLEVARVESFHGFIYGCPEQAPRLMDYLGDAWYLEPFFKHSGL 180
Db      119 KELYGALDKKCKGKLEVARVESFHGFIYGCPEQAPRLMDYLGDAWYLEPFFKHSGL 178
QY      181 ELVGPCKVIVIKANKKAPAEVFGDAYHVGMTHASSLSRSGESIFSSLAGNAALPREGAGL 240
Db      179 ELVGPCKVIVIKANKKAPAEVFGDAYHVGMTHASSLSRSGESIFSSLAGNAALPREGAGL 238
QY      241 QMTSKYSGMGVLMDSYGVSHADYLPBELMAFGAKOEKRLNEIGVRRARIYRSHLNCYV 300
Db      239 QMTSKYSGMGVLMDSYGVSHADYLPBELMAFGAKOEKRLNEIGVRRARIYRSHLNCYV 298
QY      301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRRLADSVORTTGPAGFWS 360
Db      299 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRRLADSVORTTGPAGFWS 358
QY      361 DDNDNMETASONGKATYQSDSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 420
Db      359 DDNDNMETASONGKATYQSDSDLSNLGFGEDVYGDVYPGVYKSAIGETSYRGFYRAY 418
QY      421 QAHVSSNMAEFEBHASTWHTTELTKTTDR 449
Db      419 QAHVSSNMAEFEBHASTWHTTELTKTTDR 447

```

```

RESULT 5
O52382      PRELIMINARY:      PRT:      447 AA.
AC      O52382:      01-JUN-1998 (TREMBlrel. 06, Created)
DT      01-JUN-1998 (TREMBlrel. 06, last sequence update)
DT      01-DEC-2001 (TREMBlrel. 19, last annotation update)
DE      Naphthalene 1,2 dioxygenase large oxygenase component.
GN      NAGAC.
OS      Ralstonia sp. U2.
OC      Plasmid pMWU2.
OX      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
RN      NCBI_TaxID=70356;
RP      [1]
RC      SEQUENCE FROM N.A.
RX      MEDLINE=98233751; PubMed=9573207;
RA      Fuemmayer S.L., Wild M., Boys A.U., Williams P.A.;
RT      "A gene cluster encoding steps in conversion of naphthalene to
      gentisate in Pseudomonas sp. strain U2."
RL      J. Bacteriol. 180:2522-2530(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=U2;
RX      MEDLINE=20576173; PubMed=11133965;
RA      Zhou N.Y., Fuemmayer S.L., Williams P.A.;
RT      "nad genes of ralstonia (Formerly pseudomonas) sp. strain U2 encoding
      enzymes for gentisate catabolism."
RL      J. Bacteriol. 183:700-708(2001).
DR      EMBL; AF036940; AAD12610.1; -.
DR      HSSP; P23094; INDO.
DR      InterPro; IPR001281; Rleske.
DR      InterPro; IPR001663; Ring_hydroxyl_A.
DR      Pfam; PF00355; Rleske; 1.
DR      Pfam; PF00848; Ring_hydroxyl_A; 1.
DR      PRINTS; PR00090; RINGDIOXGNASE.
DR      PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW      Dioxygenase; Plasmid.
SQ      SEQUENCE 447 AA; 49570 MW; 4553AFA4410ED0 CRC64;

```

```

Query Match      90.8%; Score 2187; DB 2; Length 447;
Best Local Similarity 89.1%; Pred. No. 1.3e-158;
Matches 400; Conservative 25; Mismatches 22; Indels 2; Gaps 1;
QY      1 MNYNKILVSESGLSQKHLIHGDEELFQHELKTIIFARNMLFLTHDSLIPAPGDYATAKKG 60

```

Db	1	MIYEN--LYSEAGLQKRLHIGDKELFQHELKTIFARNWLELTHDSLLIPSGDYVYAKMG	58
Qy	61	IDEVIVSRQNDGSTRAFILVNCRRHKGKTLVSYEAGNAGFCWCSYHGWGFGSNGELQSVPE	120
Db	59	VDEVIVSRQNDGSTRAFILVNCRRHKGKTLVHAEGNAGFCWCSYHGWGFGSNGELQSVPE	118
Qy	121	KDLYESLNRKKCKLGLKEVARYESHGFYTCGFDDEAPPLMDYLGDAAMYLEPMFKHSGGL	180
Db	119	KELYDITIKKKCKLGLKEVRIESEHGFYTCGFDDEAPPLMDYLGDAAMYLEPIFKHSGGL	178
Qy	181	ELVGPBGVVIRKANKAAEKNVGDAYVIGTTHASSLRSGESTSSLAGNAALPREBAGL	240
Db	179	ELVGPBGVVIRKANKAAEKNVGDAYVIGTTHASSLSRSGSITPLAGNAMPREBAGL	238
Qy	241	QMTSKYSGGMGLMDGYESGVSHADLYPELMAFGGAKOERLKEIGDVARIRYSHLNCIV	300
Db	239	QMTSKYSGGMGLMDGYESGVSHADLYPELMAFGGAKOERLKEIGDVARIRYSHLNCIV	298
Qy	301	FPNNSMLTCSGVFKVWNPIDANTTEVWMTYIAVEKMDPEDLKRLRADSVORTGPAGFWS	360
Db	299	FPNNSMLTCSGVFKVWNPIDANTTEVWMTYIAVEKMDPEDLKRLRADSVORTGPAGFWS	358
Qy	361	DDNDMEFASONGKKYSGRSDLSNLGCFGEDYVGDAYPCGVGKSAIGETSTRGFTYRAY	420
Db	359	DDNDMEFESQAKKYQSSNSDLIANLGFGKDYGVGDECEYPCGVAKSAIGETSYRGFTYRAY	418
Qy	421	QAHVSSSWMAFEHASSYWHETELKRTDTR	449
Db	419	QAHSSSWMAFEENTSRMWHETELKRTDTR	447

Db	1	MSYON -- LVSSEALQTHKLITLYDCKELPQHELIKTI	FAFNHMLFELHDSLIPSRODYIKAKMG	58	
Qy	61	IDEVIVSRONDGSI	RAFLVWCNHRKCTLVSYVEAGNAKGFVCSYHGMFGSNGELOSVPE	120	
Db	59	VDEVIVSRONDGSV	RAFLVWCNHRKCTIYVAEAGNAKGFVCSYHGMFGSNGELOSVPE	118	
Qy	121	KDLTGESLNRKCLGL	KEVARDESHGFIYGCDFODAPRLMDIYIGDAWYLEPMRKHSGL	180	
Db	119	KELTGDIAIKKCKGL	KEVPRIESEFHGFIYGCDFODAPRLDIYIGDAWYLEPTEKHSGL	178	
Qy	181	ELVGPCKVIVYKANK	KAPENVGDAIVGVTGTHASLRSESESTFSLGAMALPREAGL	240	
Db	179	ELVGPCKVIVYKANK	PLAENVGDAIVGVTGTHASLIRAGQSTIFAPLAGNAAPREAGL	238	
Qy	241	QMTSKYSGSGV	LMDYSCGVSHADYVPELMAFGAKOERLNKIRGVARIRYSHLNTV	300	
Db	239	QMTTKYSGIGV	LMDYSCIGSADWYVPEMAFGAKOEKLAKEIGDYRARIYRSLNGTV	298	
Qy	301	FPNNSMLTCSGV	FYKWNPIDANTTEVMTYAI	VERKDEMEDLRRLADSVORTGPAGWES	360
Db	299	FPNNSFLTCSGV	FYKWNPIDENTTEFVMTYAI	VERKDEMEDLRRLADVAORSGPAGWES	358
Qy	361	DDNDNMFASQNGK	KYQSDLSNLGFGEDVYGDAIVPGVYKSAIGETSYRGFTYAY	420	
Db	359	DDNDNMGTLISQNAK	KYQSSNSDLADLFGKDVYDECPYGVYKSAISETSYRGFTYAY	418	
Qy	421	QAHVSSSNAAEP	EBHASTYHTELTKTDR	449	
Db	419	QAHSSSNAAEP	EBNTSRMHTELTKTDR	447	

ID	Q8VUDA	PRELIMINARY;	PRT;	447 AA.
AC	Q8VUDA			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	DnaC.			
OS	Burkholderia cepacia (Pseudomonas cepacia).			
OC	Bacteria; Proteobacteria; Beta subdivision; Burkholderia group;			
OC	Burkholderia.			
OX	NCBI_TaxID=292;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-R34;			
RX	MEDLINE=20254695; PubMed=10795678;			
RA	Johnson G.R., Jain R.K., Spain J.C.;			
RT	"Properties of the trihydroxytoluene oxygenase from Burkholderia			
RT	cepacia R34: an extradiol dioxygenase from the 2,4-dinitrotoluene			
RT	pathway.";			
RL	Arch. Microbiol. 173:86-90(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-R34;			
RA	Johnson G.R., Jain R.K., Spain J.C.;			
RT	"Origins of the 2,4-dinitrotoluene pathway.";			
RL	Submitted (JUN-2001) to the EMBL/genbank/DBD databases.			
DR	EMBL; AFI69302; AAL50021.1; -			
DR	InterPro; IPR001281; Rtske.			
DR	InterPro; IPR001663; Rtn_hydroxyl_A.			
DR	Pfam; PF003355; Rtske; 1.			
DR	Pfam; PF00848; Rtn_hydroxyl_A; 1.			
DR	PRINTS; PRO0090; RINGDIOXGNSE.			
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.			
SO	SEQUENCE 447 AA; 49468 MW; 2A5BF8558320275C CRC64;			
Query Match	86.0%; Score 2071; DB 2; Length 447;			
Best Local Similarity	83.5%; Pred. NO.1.le-150;			
Matches	375; Conservative 38; Mismatches 34; Indels 2; Gaps			

RESULT	7
P95564	PRELIMINARY; PRT; 447 AA.
AC	P95564;
DT	01-MAY-1997 (TrEMBLrel. 03, Created)
DT	01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	ISpA1pha 2NT.
CN	NPfAC.
OS	Pseudomonas sp.
OC	Bacteria; Proteobacteria.
OX	NCBI_TaxID=306;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-JS42;
RX	MEDLINE=97128768; PubMed=8973308;
RA	Parales J.V., Kumar A., Parales R.E., Gibson D.T.;
RT	"Cloning and sequencing of the genes encoding 2-nitrotoluene
RT	dioxygenase from Pseudomonas sp. JS42.";
RL	Gene 181:57-61(1996).
DR	EMBL: U49504; AABA0383.1; -
DR	HSP; P23094; INDO.
DR	InterPro: IPR001281; RlsAke.
DR	InterPro: IPR001663; Ring_hydroxyl_A.
DR	Pfam: PF00355; RlsAke; 1.
DR	Pfam: PF00848; Ring_hydroxyl_A. 1.
DR	PRINTS: PR00090; RINGDIOXGNASE.
DR	PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
SQ	SEQUENCE 447 AA; 49485 MW; 1CB0E223E528E3BD CRC64;
QY	Query Match 85.1%; Score 2051; DB 2; Length 447;
	Best Local Similarity 83.7%; Pred. No. 3,7e-149;
	Matches 376; Conservative 34; Mismatches 37; Indels 2; Gaps 1.
Db	1 MSYGN-LVSEAGLTQKLLHGDGKELFQHELKTIFARWMLFLTHDSLIPSGDYVAKMG 58
QY	1 MNYNKLIVSESGLSQKHLHGDEELFQHELKTIFARNMFLFTHDSLIPAGDYVAKMG 60
Db	1 MSYGN-LVSEAGLTQKLLHGDGKELFQHELKTIFARWMLFLTHDSLIPSGDYVAKMG 58
QY	61 IDEYIVSRQDGSIRAFVLANVCRHKGTLIVSYAGNAKGFVCSHYGMFGSGNGELQSVPE 120
Db	59 VDEYIVSRQDGSIRAFVLANVCRHKGTLIVSYAGNAKGFVCSHYGMFGSGNGELQSVPE 118

QY	121	KDLGESLNNKCCJCLCEKAVARESHHGTYTGCFOEADPLMDITGDAAWYLEPFKHSGGI	180
Dd	119	KELYGDAKKCKCJLCEKVPRIESEHGFYICGCPAEAPLPIDYLGDAWYLEPFFKHSGGI	178
QY	181	ELYBPGGVVYKAWMKPAEFNFVDAAHYVOTHTAASSLRSESTFSSLAGNAALPREGAGI	240
Dd	179	ELYBPGGVVYKAWMKPAEFNFVGDIYHVGTTHAALLRAQSOVSFSSLAGNAKLPREGAGI	238
QY	241	QMTSKYSGCMVLVDGYSGVSIADLYPELMAFGAKOERLNKEIGDVARITYRSHLNCITY	300
Dd	239	QMTSKYSGCMGLTDYYISGNFSADMVFDLMAFGAKOEKLAKEIGDVARITYRSILNGTV	298
QY	301	FPNNSMLTSCGVFWNMPIDANTTEVWTYAIVEKDMEDEKRLRADSVQRTTGPAFTWES	360
Dd	299	FPNNSFLTGSAITEFKWNPIIDENTTEEVWTYAFVEKDMEDEKRLRLADAQRSIGTPAFTWES	358
QY	361	DDNDNMETASONGKRYOSRSDLSLNLGFGEDYVYGAUVYPGVYGSKAIGETSYRGFYRAY	420
Dd	359	DDNEEMETLSONAKKYOSSNSDOIASLGFGKDYVYGEDCYGVYGSKAIGETSYRGFYRAY	418
QY	421	QAHHSSNWAEFEHASSTWHTELTKTTDR	449
Dd	419	QAHHSSNWAEFEENASRNWHTELTKTTDR	447

RESULT	ID	PRELIMINARY:	PRT:	447	AA.
08RTL4	08RTL4				
AC	08RTL4;				
DT	01-JUN-2002 (TrEMBLrel. 21, Created)				
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)				
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)				
DE	Oxygenase-alpha NBD0.				
GN	NB2AC.				
OS	Comamonas sp. JS765.				
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.				
OX	NCBI_TaxID=58226;				
	[1]				
RP	SEQUENCE FROM N.A.				
RP	STRAIN=JS765;				
RC	MEDLINE=21681044; PubMed=11823201;				
RX	Lesener D.J., Johnson G.R., Patales R.E., Spain J.C., Gibson D.T.;				
RT	"Molecular Characterization and Substrate Specificity of Nitrobenzene				
RT	Dioxygenase from Comamonas sp. Strain JS765.";				
RL	Appl. Environ. Microbiol. 68:634-641(2002).				
DR	EMBL; AF379638; AAL76202.1; -				
SO	SEQUENCE 447 AA; 49556 MW; 4CA96A1F99A37DF3A CRC64;				

Query Match	83.9%	Score 2022;	DB 2;	Length 447;
Best Local Similarity	82.0%;	Pred. No. 6.2e-147;		
Matches 368;	Conservative 40;	Mismatches 39;	Indels 2;	Gaps 1;

[illegible][illegible]

RESULT	9
ID	Q45695
AC	Q45695;
DT	01-NOV-1996 (TREMBLrel_01, Created)
DT	01-NOV-1996 (TREMBLrel_01, Last sequence update)
DT	01-DEC-2001 (TREMBLrel_19, Last annotation update)
DE	ISP-alpha.
GN	DNTAC.
OS	Burkholderia sp. (strain RASC).
OC	Bacteria; Proteobacteria; beta subdivision; Burkholderia group; Burkholderia.
OX	NCBI_taxid=69003;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=DNT;
RX	MEDLINE=96326338; PubMed=8759857;
RA	Suen W.C., Haigler B.E., Spain J.C.;
RT	"2,4-Dinitrochloroethene dioxygenase from Burkholderia sp. strain DNT:
RL	similarity to naphthalene dioxygenase.";
J	J. Bacteriol. 178:4926-4934(1996).
DR	EMBL; U62430; AAB09766.1; -;
DR	HSSP; P23094; INDO.
DR	InterPro; IPR001281; Rleske.
DR	InterPro; IPR001663; Ring_hydroxyl_A.
DR	Pfam; PF00355; Rleske; 1.
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.
DR	PRINTS; PRO0090; RNGLDIOXGNASE.
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
SEQ	SEQUENCE 451 AA; 49827 MW; 726796C29CE9A10 CRC64;

Query Match	81.9%	Score 1973.5;	DB 2;	Length 451;
Best Local Similarity	80.2%	Pred. No. 3,3e-143;		
Matches 360;	Conservative 41;	Mismatches 45;	Indels 3;	Gaps 2

[illegible]

||||| ||| ||||| :|||:||||| ||||| ||| :|||:|||||  
Db 364 DDNDNM-VLSQNAKKYQSSNSDLADLGFQKDYGDCEYGVVSKSAFSETHRGFYRAY 422  
QY 421 QAHVSSNAAEFEHASTHTLTKTTDR 449  
|||:|||||:| |||||  
Db 423 QAHISSNAAEFEHASTHTLTKTTDR 451

## RESULT 10

Q9ETK2 PRELIMINARY; PRT; 277 AA.  
AC Q9ETK2;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).  
GN NAHAC2 OR NAHAC.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=21DINH, PRIM1, AND 31A2NH;  
RA Ferrero M.A., Lalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";  
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF306436; AAC25698.1; -;  
DR EMBL: AF306432; AAC25692.1; -;  
DR EMBL: AF306436; AAC25696.1; -;  
DR HSP: P23094; INDO.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIPOXNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
FT NON\_TER 1 277  
FT NON\_TER 277 1  
SQ SEQUENCE 277 AA; 30094 MW; BFCF9B5A379546DB CRC64;

## Query Match 61.9%; Score 1490; DB 2; Length 277;

Best Local Similarity 100.0%; Pred. No. 2.1e-106;  
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRQNDGSIKRAFLNVCRRHCKTLVSVAGNAKGFVCSYHGMGFGSGNGELO 115  
|||:|||||:| |||||  
Db 1 TAKMGIDEVIVSRQNDGSIKRAFLNVCRRHCKTLVSVAGNAKGFVCSYHGMGFGSGNGELO 60  
QY 116 SVPEKDLVGSLSNKKCLGLKEVARVESFHGFTYGCDFDQEPAPLMDYLGDAWYLEPMEK 175  
|||:|||||:| |||||  
Db 61 SVPEKDLVGSLSNKKCLGLKEVARVESFHGFTYGCDFDQEPAPLMDYLGDAWYLEPMEK 120  
QY 176 HSGGLELVGPPGKVYIKANKKAPAEFVGDYHVGWTHASSLSRSGSIFSSLAGNAALPP 235  
|||:|||||:| |||||  
Db 121 HSGGLELVGPPGKVYIKANKKAPAEFVGDYHVGWTHASSLSRSGSIFSSLAGNAALPP 180  
QY 236 EGAGLQMTSKYSGMGVLMGDSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 295  
|||:|||||:| |||||  
Db 181 EGAGLQMTSKYSGMGVLMGDSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 240  
QY 296 LNCIVFPNNSMLTCSGVFKVWNPIDANTTEWMTYAIY 332  
|||:|||||:| |||||  
Db 241 LNCIVFPNNSMLTCSGVFKVWNPIDANTTEWMTYAIY 277

## RESULT 11

Q9F5S4 PRELIMINARY; PRT; 277 AA.  
AC Q9F5S4;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).  
GN NAHAC2.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=51TINASAL;  
RA Ferrero M.A., Lalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";  
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF306440; AAC25700.1; -;  
DR HSP: P23094; INDO.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIPOXNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
FT NON\_TER 1 277  
FT NON\_TER 277 1  
SQ SEQUENCE 277 AA; 30094 MW; FADFDA5D37E241AB CRC64;

## Query Match 61.8%; Score 1488; DB 2; Length 277;

Best Local Similarity 99.6%; Pred. No. 3e-106;  
Matches 276; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRQNDGSIKRAFLNVCRRHCKTLVSVAGNAKGFVCSYHGMGFGSGNGELO 115  
|||:|||||:| |||||  
Db 1 TAKMGIDEVIVSRQNDGSIKRAFLNVCRRHCKTLVSVAGNAKGFVCSYHGMGFGSGNGELO 60  
QY 116 SVPEKDLVGSLSNKKCLGLKEVARVESFHGFTYGCDFDQEPAPLMDYLGDAWYLEPMEK 175  
|||:|||||:| |||||  
Db 61 SVPEKDLVGSLSNKKCLGLKEVARVESFHGFTYGCDFDQEPAPLMDYLGDAWYLEPMEK 120  
QY 176 HSGGLELVGPPGKVYIKANKKAPAEFVGDYHVGWTHASSLSRSGSIFSSLAGNAALPP 235  
|||:|||||:| |||||  
Db 121 HSGGLELVGPPGKVYIKANKKAPAEFVGDYHVGWTHASSLSRSGSIFSSLAGNAALPP 180  
QY 236 EGAGLQMTSKYSGMGVLMGDSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 295  
|||:|||||:| |||||  
Db 181 EGAGLQMTSKYSGMGVLMGDSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 240  
QY 296 LNCIVFPNNSMLTCSGVFKVWNPIDANTTEWMTYAIY 332  
|||:|||||:| |||||  
Db 241 LNCIVFPNNSMLTCSGVFKVWNPIDANTTEWMTYAIY 277

## RESULT 12

Q9F5T3 PRELIMINARY; PRT; 277 AA.  
AC Q9F5T3;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).  
GN NAHAC.  
OS Pseudomonas sp. 191INDH.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=139738;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=191INDH;  
RA Ferrero M.A., Lalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";  
RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

OV 236 EGAGTOMTSKYSGMGV.LWDGVSQVHSADL.VDET.MAFGAKOEPT.NKETGDPAPTYRSH 295

DB 181 EGAGLOMTSKYSGMGVLMGYSVHSADLVPELMARFGAKQEGLNKEIGDVRARIYRSH 240  
QY 296 LNCTVFPNNSMLTCSGVFKVWNPIDANTEWMTYAIV 332  
DB 241 LNCTVFPNNSMLTCSGVFKVWNPIDANTEWMTYAIV 277

## RESULT 15

09F553 PRELIMINARY; PRT; 277 AA.  
ID 09F553  
AC 09F553;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 1 (Fragment).  
GN NAHAC1.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=511ANH;  
RA Ferrero M.A., Lalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in  
Pseudomonas strains from West Mediterranean Sea";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF306441; ANG25701.1; -.  
DR HSSP; P23094; INDO.  
DR InterPro; IPR001281; Rieske.  
DR InterPro; IPR001653; Ring\_hydroxyl\_A.  
DR Pfam; PF00355; Rieske; 1.  
DR PRINTS; PR000848; Ring\_hydroxyl\_A; 1.  
DR PROSITE; PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
FT NON\_TER 1  
FT NON\_TER 277  
SQ SEQUENCE 277 AA; 30062 MW; 82C78D865A92D0AB CRC64;

## Query Match

Best Local Similarity 60.2%; Score 1451; DB 2; Length 277;

Matches 268; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRONDGSTRAPLVNCRHRGKTLVSVAGNAKGFVCSYHGMFGSGNELQ 115  
DB 1 TAKMGIDEVIVSRQSDGSTRAPLVNCRHRGKTLVNAEAGNAKGFVCSYHGMFGSGNELQ 60  
QY 116 SVPEKDLXGESLNKKCLGLKEVARVESFHGFLYGCFOEAPPLMDYLGDAWYLEPMFK 175  
DB 61 SVPEKELXGESLNKKCLGLKEVARVESFHGFLYGCFOEAPPLMDYLGDAWYLEPMFK 120  
QY 176 HSGGLELVGPGRKVIKANKKAPAEENVGDAYHYVWTHASSLSRSGESIFSSLAGNAALPP 235  
DB 121 HSGGLELVGPGRKVIKANKKAPAEENVGDAYHYVWTHASSLSRSGESIFSSLAGNAALPP 180  
QY 236 EGAGLOMTSKYSGMGVLMGYSVHSADLVPELMARFGAKQERLNKEIGDVRARIYRSH 295  
DB 181 EGAGLOMTSKYSGMGVLMGYSVHSADLVPELMARFGAKQERLNKEIGDVRARIYRSH 240  
QY 296 LNCTVFPNNSMLTCSGVFKVWNPIDANTEWMTYAIV 332  
DB 241 LNCTVFPNNSMLTCSGVFKVWNPIDANTEWMTYAIV 277

Search completed: January 27, 2003, 09:02:22  
Job time : 27.8333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:47:32 : Search time 7.16667 Seconds  
(without alignments)  
2598.540 Million cell updates/sec

Title: US-09-843-250-34  
Perfect score: 2409  
Sequence: 1 MNYNKILVSESGLSQKHLI.....AEFHASSTWHTTELTQTD 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2402	99.7	449	1 NDOB_PSEPU	P23094 pseudomonas
2	2316	96.1	449	1 NDOB_PSEFL	O07824 pseudomonas
3	2217	92.0	449	1 NDOB_PSEAB	Q51494 pseudomonas
4	735	30.5	453	1 HCAE_ECOLI	Q47139 escherichia
5	650.5	27.0	450	1 BEDL_PSEPU	O07944 pseudomonas
6	648.5	26.9	450	1 TODL_PSEPU	P13450 pseudomonas
7	615.5	25.6	458	1 BPHA_BURCE	P37333 burkholderia
8	615	25.5	457	1 BPHA_COMTE	Q46372 comamonas t
9	615	25.5	458	1 BPHA_PSEPS	P08084 pseudomonas
10	593.5	24.6	448	1 BNZA_PSEPU	O52028 pseudomonas
11	585	24.3	458	1 BNZA_PSESI	P08084 pseudomonas
12	410.5	17.0	454	1 XYIX_PSEPU	P07769 pseudomonas
13	397	16.5	461	1 BENA_PSEPU	P07769 pseudomonas
14	246	10.2	374	1 YEAW_ECOLI	P76253 acinetobact
15	239.5	9.9	442	1 CHMO_AMATR	Q93xel amarantithus
16	233	9.7	438	1 CHMO_ATRHO	O93xel amarantithus
17	224	9.3	446	1 CHMO_BETVU	O22555 beta vulgar
18	223	9.3	439	1 CHMO_SPIOL	O04121 spinacia ol
19	216	9.0	422	1 CHMO_KRATH	O96270 arabidopsis
20	165	6.8	439	1 PTH3_PSEPU	O05183 pseudomonas
21	118.5	4.9	354	1 VANA_PSEPU	O05616 pseudomonas
22	112	4.6	432	1 CBNA_COMTE	O44256 comamonas t
23	110	4.6	329	1 VANA_PSES9	P13609 pseudomonas
24	108	4.5	1276	1 PMP6_CHLNP	O94899 chlamydia p
25	104	4.3	409	1 POBA_PSEPS	O52185 pseudomonas
26	102.5	4.3	3255	1 POLG_LMVE	P89876 1 genome po
27	101.5	4.2	1411	1 Y297_HUMAN	O15040 homo sapien
28	100.5	4.2	543	1 TUP1_DROME	P16376 drosophila
29	100.5	4.2	746	1 TUP2_DROME	P16376 drosophila
30	98.5	4.1	1926	1 LPH_RABIT	P09849 corynebact
31	98.5	4.1	3255	1 POLG_LMVO	P33999 1 genome po
32	97.5	4.0	428	1 GBAL_CANAL	P28868 candida alb
33	97.5	4.0	468	1 PPAL_PICPA	P52291 plichia past

34	97	4.0	405	1 DCP2_PEA	P51851 plasm sativ
35	94.5	3.9	331	1 LDHA_RHIDE	O9pw58 rhizophila
36	94.5	3.9	350	1 UNRI_HUMAN	O9y3f4 homo sapien
37	94.5	3.9	452	1 F26_YEAST	P33604 saccharomyc
38	94.5	3.9	518	1 ATPA_ENTHR	P26679 enterococcu
39	94	3.9	847	1 ORP8_HUMAN	O9p2f1 homo sapien
40	93	3.9	420	1 DHE3_PYROO	O59650 pyrococcus
41	92.5	3.8	525	1 NCAP_MEASA	P35972 measles vir
42	92.5	3.8	1829	1 RMI_MOUSE	O70472 mus musculu
43	92	3.8	985	1 AGUL_ASPOR	O12558 aspergillus
44	92	3.8	1693	1 POLN_HEVM	O04610 hepatitis e
45	91.5	3.8	566	1 PPE_YEAST	P11491 saccharomyc

## ALIGNMENTS

RESULT 1  
ID NDOB\_PSEPU STANDARD; PRT; 449 AA.  
AC P23094: 052124: 03461: 007830:  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene  
1,2-dioxygenase ISP alpha).  
GN NDOB OR NAHAC OR DOXB OR NAH3 OR NDCC2.  
OS Pseudomonas putida, and  
OS Pseudomonas sp. (strain C18).  
OG Plasmid pDGL, Plasmid NAH7, and Plasmid NPL1.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303, 306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=89211973; PubMed=324348;  
RA Kurkela S., Lehtvaeslahti H., Palva E.T., Teeri T.H.;  
RT "Cloning, nucleotide sequence and characterization of genes encoding  
naphthalene dioxygenase of Pseudomonas putida strain NCIB9816.";  
RT Gene 73:355-362(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX Parles J.V., Parales R.E., Kumar A., Gibson D.T.;  
RA Submitted (Mar-1996) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C18;  
RX MEDLINE=94042852; PubMed=8226631;  
RA Denome S.A., Stanley D.C., Olson E.S., Young K.D.;  
RT "Metabolism of dibenzothioephene and naphthalene in Pseudomonas  
strains: complete DNA sequence of an upper naphthalene catabolic  
pathway.";  
RT J. Bacteriol. 175:6890-6901(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=G7 / ATCC 17485; PLASMID=NAH7;  
RX MEDLINE=93252277; PubMed=8486285;  
RA Simon M.J., Ossling T.D., Saunders R., Ensley B.D., Suggs S.,  
RA Harcourt A.A., Suen W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;  
RT "Sequences of genes encoding naphthalene dioxygenase in Pseudomonas  
putida strains G7 and NCIB 9816-4.";  
RT Gene 127:31-37(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=BS202; PLASMID=NPL1;  
RA Bezorodnikov S.G., Boronin A.M., Tiedje J.M.;  
RT "Nucleotide sequences of genes encoding an upper pathway of  
naphthalene metabolism of NPL1 plasmid from Pseudomonas putida strain  
BS202.";  
RT submitted (Jul-1997) to the EMBL/Genbank/DBJ databases.  
RN [6]

RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=ATCC 17484;  
RA Hamann C.;  
RU Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=98298434; PubMed=9634695;  
RA Kauppi B., Lee K., Carcedano E., Perales R.E., Gibson D.T., Eklund H.,  
Ramawamy S.;  
RT "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene  
1,2-dioxygenase";  
RL Structure 6:571-586 (1998).  
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
NAPHTHALENE DIHYDRODIOL.  
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-  
dihydronaphthalene-1,2-diol + NAD(+).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES  
CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF  
DIBENZOFIOPHENE (DBI) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO  
1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBI IS LIMITED TO  
OXIDATION OF THE AROMATIC RING.  
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME  
SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON  
SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED  
OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS  
COMPOSED OF AN HEXAMER OF THREE LARGE ALPHA SUBUNITS (NDOB) AND  
THREE SMALL BETA SUBUNITS (NDOC).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
DIOXYGENASE ALPHA SUBUNIT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M23914; AAB47591.1; -;  
DR EMBL: U49496; AAA92141.1; -;  
DR EMBL: M60405; AAA16125.1; -;  
DR EMBL: M63949; AAA25902.1; -;  
DR EMBL: AF00471; AAB62707.1; -;  
DR EMBL: AF004284; AAB61373.1; -;  
DR PIR: J50071; J50071.  
DR PIR: B49343; B49343.  
DR PDB: INDO; 23-MAR-99.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR pfam: PF00355; Rieske; 1.  
DR pfam: PF00648; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIOXGNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
KM Dioxygenase; NAD; Plasmid; 3D-structure.  
FT METAL 81 81 IRON-SULFUR (2FE-2S).  
FT METAL 83 83 IRON-SULFUR (2FE-2S).  
FT METAL 101 101 IRON-SULFUR (2FE-2S).  
FT METAL 104 104 IRON-SULFUR (2FE-2S).  
FT METAL 208 208 IRON.  
FT METAL 213 213 IRON.  
FT METAL 362 362 IRON.  
FT VARIANT 4 4 N -> K (IN STRAIN G7).  
FT VARIANT 12 12 S -> F (IN STRAIN ATCC 17484).  
FT VARIANT 15 15 S -> T (IN STRAIN G7).  
FT VARIANT 32 32 K -> R (IN STRAIN G7).  
FT VARIANT 50 50 A -> S (IN STRAIN G7).  
FT VARIANT 70 70 N -> S (IN STRAIN G7).  
FT VARIANT 90 91 SV -> NA (IN STRAIN G7).

FT VARIANT 122 122 D -> E (IN STRAIN G7).  
FT VARIANT 173 173 M -> I (IN STRAIN G7).  
FT VARIANT 225 225 S -> A (IN STRAIN G7).  
FT VARIANT 225 225 S -> C (IN STRAIN BS202).  
FT VARIANT 232 232 A -> V (IN STRAIN G7).  
FT VARIANT 275 275 A -> S (IN STRAIN G7).  
FT VARIANT 391 391 E -> K (IN STRAIN G7).  
FT VARIANT 421 421 Q -> R (IN STRAIN ATCC 17484).  
FT VARIANT 434 434 H -> D (IN STRAIN G7).  
SO SEQUENCE 449 AA; 49607 MW; 1FD2F4229664F7A8 CRC64;  
Query Match 99.7%; Score 2402; DB 1; Length 449;  
Best Local Similarity 99.8%; Pred. No. 3e-181;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNYNNKILIVSSGSLQKHILHGDDELFOHLEKTFANWLFTHDSLIPAGDYVAKMG 60  
DB 1 MNYNNKILIVSSGSLQKHILHGDDELFOHLEKTFANWLFTHDSLIPAGDYVAKMG 60  
QY 61 IDEVIVSRONDGSIKAFILNVCRRHGRKTLVSVEAGNAKGFVCSYHGNGFSGNELQSVPE 120  
DB 61 IDEVIVSRONDGSIKAFILNVCRRHGRKTLVSVEAGNAKGFVCSYHGNGFSGNELQSVPE 120  
QY 121 KDLYGESLKKCLGLKEVARYESFHGTYGCPDQEPPLMDYIGDAAWYLEPFRKSGGL 180  
DB 121 KDLYGESLKKCLGLKEVARYESFHGTYGCPDQEPPLMDYIGDAAWYLEPFRKSGGL 180  
QY 181 ELVPPGKVVYKAMKPAENFVGDADAYHGWTHASSLRSGSEIFSSLAGNAALPPRGAGI 240  
DB 181 ELVPPGKVVYKAMKPAENFVGDADAYHGWTHASSLRSGSEIFSSLAGNAALPPRGAGI 240  
QY 241 QMTSKYSGMGVLWDGYSGVHSDLVPELMAFGAKOERLKEIGDVARIRYRSHLNCIV 300  
DB 241 QMTSKYSGMGVLWDGYSGVHSDLVPELMAFGAKOERLKEIGDVARIRYRSHLNCIV 300  
QY 301 FPNNSMLTCSGVFWNMPIDANTTEVNTYAIVERKDPEDLKRRLASVQRTTPAPFWS 360  
DB 301 FPNNSMLTCSGVFWNMPIDANTTEVNTYAIVERKDPEDLKRRLASVQRTTPAPFWS 360  
QY 361 DDNNMETASONGKRYOSRSDLSNLGFGEVDYGDVAYGVVKGKSAIGETSYRGFRAY 420  
DB 361 DDNNMETASONGKRYOSRSDLSNLGFGEVDYGDVAYGVVKGKSAIGETSYRGFRAY 420  
QY 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449  
DB 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449  
RESULT 2  
NDOB\_PSEFL STANDARD; PRT; 449 AA.  
AC 007824;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Naphthalene 1,2-dioxygenase alpha subunit (Ec 1.14.12.12) (Naphthalene  
1,2-dioxygenase ISP alpha).  
GN NDOB OR NDOC2.  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 17483;  
RA Hamann C.;  
RU Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
NAPHTHALENE DIHYDRODIOL.  
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-



```
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF004283; AAB61370.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske.1.
DR Pfam: PF00848; Ring_hydroxyl_A.1.
DR PRINTS: PS00570; RING_HYDROXYL_ALPHA.1.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA.1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
DR Dioxygenase; NAD; Plasmid.
FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 208 208 IRON (BY SIMILARITY).
FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 362 362 IRON (BY SIMILARITY).
SQ SEQUENCE 449 AA; 49512 MW; 186A7924AD437A1C CRC64;

Query Match 96.1%; Score 2316; DB 1; Length 449;
Best local similarity 95.5%; Pred. No. 1.7e-174;
Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 MNNKILVSESGLTQKHILHGDELFQHELTIFARNMLFTHDSLIPAGDYVAKKG 60
DB 1 MNNKILVSESGLTQKHILHGDELFQHELTIFARNMLFTHDSLIPAGDYVAKKG 60
QY 61 IDEVIYSRONDSIRAFILVNCVCRHGRKTIYSVEGNNAKGFVCSYHGSGFSGELQSVPE 120
DB 61 IDEVIYSRONDSIRAFILVNCVCRHGRKTIYSVEGNNAKGFVCSYHGSGFSGELQSVPE 120
QY 61 IDEVIYSRONDSIRAFILVNCVCRHGRKTIYSVEGNNAKGFVCSYHGSGFSGELQSVPE 120
DB 61 IDEVIYSRONDSIRAFILVNCVCRHGRKTIYSVEGNNAKGFVCSYHGSGFSGELQSVPE 120
QY 121 KDLGSESLNKKCLGLKEVAVRESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180
DB 121 KDLGSESLNKKCLGLKEVAVRESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180
QY 121 KDLGSESLNKKCLGLKEVAVRESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180
DB 121 KDLGSESLNKKCLGLKEVAVRESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180
QY 181 ELVGPGRKVIYKANKAPAEVGYDAIYHVGWTHASSLRGSGSTFSSLAGNALPPGAGI 240
DB 181 ELVGPGRKVIYKANKAPAEVGYDAIYHVGWTHASSLRGSGSTFSSLAGNALPPGAGI 240
QY 181 ELVGPGRKVIYKANKAPAEVGYDAIYHVGWTHASSLRGSGSTFSSLAGNALPPGAGI 240
DB 181 ELVGPGRKVIYKANKAPAEVGYDAIYHVGWTHASSLRGSGSTFSSLAGNALPPGAGI 240
QY 241 QMTSKGSGMGVLMDSYGVSHADVLPELMAFGAKOEKRLNKEIGVRAIRYSHLNCIV 300
DB 241 QMTSKGSGMGVLMDSYGVSHADVLPELMAFGAKOEKRLNKEIGVRAIRYSHLNCIV 300
QY 241 QMTSKGSGMGVLMDSYGVSHADVLPELMAFGAKOEKRLNKEIGVRAIRYSHLNCIV 300
DB 241 QMTSKGSGMGVLMDSYGVSHADVLPELMAFGAKOEKRLNKEIGVRAIRYSHLNCIV 300
QY 301 FPNNSMLTSGSGVFKVWNPIDANTTEVYTAIYEKDPEDLKRRLADSVORTTGPAGFWS 360
DB 301 FPNNSMLTSGSGVFKVWNPIDANTTEVYTAIYEKDPEDLKRRLADSVORTTGPAGFWS 360
QY 301 FPNNSMLTSGSGVFKVWNPIDANTTEVYTAIYEKDPEDLKRRLADSVORTTGPAGFWS 360
DB 301 FPNNSMLTSGSGVFKVWNPIDANTTEVYTAIYEKDPEDLKRRLADSVORTTGPAGFWS 360
QY 361 DDNDNMETASQNKRYQSRDLSLNLGFGEDYVGYDAVPGVYKSAIGETSYRGFYRAY 420
DB 361 DDNDNMETASQNKRYQSRDLSLNLGFGEDYVGYDAVPGVYKSAIGETSYRGFYRAY 420
```

```
DB 361 DDNDNMETASQNKRYQSRDLSLNLGFGEDYVGYDAVPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNNMAEFEPDASTWHTTELKTTDR 449
DB 421 QAHVSSNNMAEFEPDASTWHTTELKTTDR 449

RESULT 3
NDOB.PSEAE STANDARD; PRT; 449 AA.
AC 051494;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene
DE 1,2-dioxygenase ISP alpha).
GN NDOB OR PAK13.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAK1;
RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,
RA Kiyohara H.;
RT "The molecular analysis of an NAH7-type gene cluster, pah, located on
RT the chromosome of Pseudomonas aeruginosa PAK1."
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL.
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D84146; BAA12240.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske.1.
DR Pfam: PF00848; Ring_hydroxyl_A.1.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA.1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
DR Dioxygenase; NAD.
FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 208 208 IRON (BY SIMILARITY).
```

FT	METAL	213	213	IRON (BY SIMILARITY).
FT	METAL	362	362	IRON (BY SIMILARITY).
SO	SEQUENCE	449 AA:	49715 MW:	35a189136722a21c CR664:
	Query Match	92.0%;	Score 2217;	DB 1; Length 449;
	Best Local Similarity	89.8%;	Pred. No. 1e-166;	
	Matches 403; Conservative	23;	Mismatches 23;	Indels 0; Gaps 0;
Qy	1	MNYNNKLIIVSESGLSCKHLIHGDEDFEHQELHTIFARNMLFLTHDSLIAPGQVYAKMG	60	
Db	1	MMYKNNLVSESGLTQKHLIHGDEDFEQLELTIFARNMLFLTHDSLIAPGQVYAKMG	60	
Qy	61	IDEVIVSRONDSSIRAFNLVWCRRKGTLLVSVEAGNAKGVCSYHMGFGSGNGLQSVPE	120	
Db	61	VDEVIVSRONDSSIRAFNLVWCRRKGTLLVHAEGNAKGVCSYHMGFGANGELQSVPE	120	
Qy	121	KDLVGEISLKKCLGLKEVARVSEFHGFTYGCDEQEAAPLMYLDGDAWTLPEMFKHSGL	180	
Db	121	KELYEELADKKCMGKEVARVSEFHGFTYGCDEEAAPLSKDYMGDGWVLEPFKHSGL	180	
Qy	181	ELVGPGRKVIKANKKAPARENEVGDYVHYGWMVHASSLRSGESIFSSLAGNAALPPRGAGL	240	
Db	181	ELIGRPGKVIKANKKAPARENTGDAHYHGWTHASSLRSGQSYFSSLAGNAALPPRGAGL	240	
Qy	241	QMTSKYSGSGMGLMDGYSGVSHADLVPELMAFGAKQERLNKEIGDVARIRYRSHLCTV	300	
Db	241	QMTSKYSGSGMGLMDGYSGVSHADLVPELMAFGAKQERLNKEIGVARIRYRSHLCTV	300	
Qy	301	PNNNSMLTCSGFYKWNPLFDANTTVEWYTAIYEKMPEDLKRRLADSVORTGPAFWMS	360	
Db	301	PNNSFLLTCSGFYKWNPLFDANTTVEWYTAIYKEMKPEDLKRRLADVAORTGPAFWMS	360	
Qy	361	DDNDMMETASQNGKRYKQSRDLSNLGFGCEVYGDVAVPYGVGKSAIETSYRGYRAY	420	
Db	361	DDNDMMETVSNQAKKYSRSDGDLVSNLGRGQVYGDVYRGVYKSAIETSYRGYRAY	420	
Qy	421	QAHVSSSNMAEFEFHASSYTHTELTETTR	449	
Db	421	GAHISSSWAEEFEDVSKNMHTLAKTTDR	449	
	RESULT 4			
	HCAE_ECOLI	STANDARD:	PRT:	453 AA.
AC	047139; P77590; P78203;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	3-phenylpropanoate dioxygenase alpha subunit (EC 1.14.1.-)			
DE	(Dioxygenin alpha subunit).			
CN	HCAE OR PHDC1 OR HCAE OR HCAAI OR DIGA OR B2538 OR Z3809 OR ECS3404.			
OS	Escherichia coli, and			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RP	SEQUENCE FROM N.A.			
RP	STRAIN=K12;			
RA	Tulin E., Gasser F., Biville F.;			
RL	submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.			
RA	STRAIN=K12 / MG1655;			
RA	MEDLINE=97426617; PubMed=9278503;			
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M., Collins-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
FT	"The complete genome sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474(1997).			
RP	SEQUENCE FROM N.A.			

[illegible]

```

FT METAL 108 108 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 218 218 IRON (BY SIMILARITY).
FT CONFLICT 20 20 V -> A (IN REF. 1).
FT CONFLICT 384 453 GHANRSKLCLEMGIQGRKRRDGIPTNTYIFSETAARG
FT YORADLSSKLSMGEQVLTAKTAAYOOEYVK -> ATAAPTAN
FT CVMKGIKRSAAATTAFLALLTISFQKPLVECTNAGPIF
FT (IN REF. 1).
FT
FT
FT
FT
SO SEQUENCE 453 AA; 51109 MW; 02535BF5F47643FD CRC64;

Query Match 30.5%; Score 735; DB 1; Length 453;
Best Local Similarity 36.4%; Pred. No. 2; 75-50;
Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps

QY 20 IHGDEELFOHELKITFARNWLEFLTHDSLIPADGDYVYAKMGIDEVIVSRQNDGSIKRAFLN 79
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 24 IYTDIDDIYOLELIERIFFGKCMLEFLAHESQIPKRGDFPNTMGDAVVVYRQKDSIKAFIN 83
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 80 VCRHHGKTLVSEANGNAGFVCSYHGMRGFSNGELOSVPFEEDLYGESLNKKGCLIKETVA 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 84 QCRHRAHMYSAVADCGNTRAFETCPYHGWSYGINGELIDVLEPRAYPDQGLCKSHMGINLEVP 143
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 140 RVSEPHGEFTYGFCEPDDEAPPLMDYLDLDAAMYLEPMF-KHSGLELYVGPCKVVIKANWKAP 198
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 144 CVESKTKGLIFGWMIDTSAAGLRDYLGDIDAMYLDMGLDRRGCGTEIYVGQVKNVINCWKKEP 203
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 AENFVGDAYVHWGWTNASSL-----RSGSEIFSLGNAALP----- 234
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 204 AEQFASDQYHALFSAVAQVYLGAKKDGSDKRLGD-----GQRAPIVWETAKDALQFG 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 235 --PEBAGIQMTSKSYSGMGVLMDDGYSYHSAIDLVEELMAFGAKOERLKEIGDVARARY 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 QDGHSGGFEETKEKPPANVW--DGAVSYSYRETYAE-----AEQRGEVYARLT 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 RSHLCCTVEFPNNNSMLTCSGVFKNVPIRANTTEVWNTYAIIVEKMDPEDKRLRLADSVQRT 352
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 304 AGHNH--IFPLLSWINGTATLTVRMPRGPDQVEVNAFCITTDKAASDEVAAEATENSATRAF 361
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 353 GPAGFWESDDNDNMTASQNGKRYOSRPSDDLNLNGFGEVDGVDAVYVPGVSKSAGTENS 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 362 GPAGFLIEDDSENMCEIQIKLKGHARRNSKLCLENGLOGEKKRRDGIPIRIT-NYIFSEFA 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 413 YRGFYRAYQAHVSSNMAEFHEHASTWHTELTK 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 421 ARGMYQWADLSSSEMOEVLDTKTAAYOOEYVK 453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
BED1_PSEPU STANDARD; PRT; 450 AA.
AC 007944;
AC 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3).
GN BEDC1.
OS Pseudomonas putida.
OG Plasmid pHM112.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RX STRAIN-ML2;
RX MEDLINE=93345820; Pubmed=8344526;
RA Tan H.-M., Tang H.-Y., Joannou C., Abdel-Wahab N.H., Mason J.R.;
RT "The Pseudomonas putida M12 Plasmid-encoded genes for benzene
RT dioxygenase are unusual in codon usage and low in G+C content.";
RL Gene 130:33-39(1993).
CC -I- CATALYTIC ACTIVITY: Benzene + NMDH + O(2) = cis-1,2-
CC dihydrobenzene-1,2-diol + NAD(+).
CC -I- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -I- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.

```

[illegible]

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Toluene 2,3-dioxygenase alpha subunit (EC 1.14.12.-).
GN TODC1.
OC Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC STRAIN=FL;
RX MEDLINE=89359301; Pubmed=2670929;
RA Zylstra G.J., Gibson D.T.;
RT "Toluene degradation by Pseudomonas putida FL. Nucleotide sequence of
   the todC1C2BAD genes and their expression in Escherichia coli.";
RL J. Biol. Chem. 264:14940-14946(1989).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: Toluene degradation; first step.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
   FERREDOXIN (TODB) AND A FERREDOXIN REDUCTASE (TODA)
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
   DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: J04996; AAA26005.1; -
DR PIR: A36516; A36516.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00370; RING-HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 222 IRON (BY SIMILARITY).
FT METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 50944 MW; 038C80F197F3485D CRC64;

Query Match 26.9%; Score 648.5; DB 1; Length 450;
Best Local Similarity 34.4%; Pred. No. 1.7e-43;
Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;

```

```

DB 264 GYVDDPMLMLAIKPKYTSWTGEPASKEAERLGSVERSKLMVEMH--TYEPTCSFL 321
OY 308 TCSCGYKWNPIDANTTEVTWYTAIVEKDPEDLKRRLADSQRTTGAGCWESDDNDNME 367
DB 322 POINTVRTWHPRGPNVEVWAFVTVADAPDDIKEEPRRQTLRTFSAGVFEODDGENWV 381
OY 368 TASQNGKKYQSRSDLSNLGFGDYGDVAYPGVYKSAIGETSYRAGFYRAYQAHVSSS 427
DB 382 EIQHTLRGKHKARSPFNKMSMDQIVNDPVYPERISNNVYSEARGLYAHLMRMKTPS 441
OY 428 NW 429
DB 442 DW 443

RESULT 7
BPHA_BURCE STANDARD; PRT; 458 AA.
ID BPHA_BURCE
AC P37333;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUN-2002 (Rel. 41, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
   dioxygenase).
GN BPHA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB400;
RX MEDLINE=92234948; Pubmed=1569021;
RA Erickson B.D., Mondello F.J.;
RT "Nucleotide sequencing and transcriptional mapping of the genes
   encoding biphenyl dioxygenase, a multicomponent
   polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain
   LB400.";
RL J. Bacteriol. 174:2903-2912(1992).
RN [2]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RC STRAIN=LB400;
RX MEDLINE=96011369; Pubmed=7592331;
RA Haddock J.D., Gibson D.T.;
RT "Purification and characterization of the oxygenase component of
   biphenyl 2,3-dioxygenase from Pseudomonas sp. strain LB400.";
RL J. Bacteriol. 177:5834-5839(1995).
RN [3]
RP ERRATUM.
RA Haddock J.D., Gibson D.T.;
RL J. Bacteriol. 178:2158-2158(1996).
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
   phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit.
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
   first step.
CC -1- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
   Bpnh subunits. A ferredoxin (Bphf) and a ferredoxin reductase
   (Bphg) must be present to obtain activity.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
   DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: M86348; AAB63425.1; -
DR PIR: B41858; B41858.

```

```
DR HSPD: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
Dioxygenase; NAD.
KM INTL MET 0
FT METAL 99 99 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 122 122 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 232 232 IRON (BY SIMILARITY).
FT METAL 238 238 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51382 MW; AF5BRL05BAFFDF4 CnC64;

Query Match 25.6%; Score 615.5; DB 1; Length 458;
Best Local Similarity 32.9%; Pred. No. 6.6e-41;
Matches 146; Conservative 69; Mismatches 186; Indels 43; Gaps 12;

QY 11 EESLQKHINHEDELPHQELKTIIPARWMLFTHSLIPAPEDDYTAKMGIDEVIVSRON 70
DB 30 EKKLLDPR-IVADQSLLELEVERGSRWMLLGHSHPETGDFLATYMGEDPPVAVRQK 88
QY 71 DGSIRAFINVCRRHKTLYSVENAGNAKGFVCSYHGMGFSGNGELQSVPEKDLGESLNK 130
DB 89 DKSIVFLNQCNRHMRICRSRAGNAKAFCTSYHGMAWDIACKLVAVPEKEAFCDKKEG 148
QY 131 KCLGLEKVE-----ARVSEFHGFIYGCEDQAPRLMDYLGDAAWYLEPMFKHS-GLLELY 183
DB 149 DC-GEDEKAEKGLQARVATYKGLVFANMDVQAPDELYLGDARPMVDMLDRTPAGTVAI 207
QY 184 GEPGKVIYKANKAKAEKPVGDAYHVG-WTHASSLRSGESIFSSLAGNAALPPEG-AGL 240
DB 208 GGMQWVWIPCNKFAAEQFCSDMYHAGTTHLSGLAG-----IPPEMLDLSQA 255
QY 241 QMTSKYSGCMGLVMDGYGVSHSADLVPELMAFGAK-----OERLNKEIGD 286
DB 256 QIPTR-GNQFRAAMGSHSGWYVDEPGSILLAVMGPKYQYWTGEPAAELAEORLHGTGMP 314
QY 287 VVARIRYRLNCTVFPNNNSMLTSGVFPVWNPIDANTEVWYTAIVERKMPEDLKRRLAD 346
DB 315 VR-RMWGQHM--TIPPTCSFLPTFNNIRIMHPRGPNIEVNAFTLVADADPAEIKEEYR 371
QY 347 SVQRTTGAGFWEESDNDNMETASONGKYSRDSLLSLNGFGEDYVGDVAVPGVVGKS 406
DB 372 HNIRNFSAGVFEODDGENWVEIQGLRGYKAKSQPLNAOMGLGRSQYTGHPDPFGNVG-Y 430
QY 407 AIGETSYRGEFYRAYOAHVSSSNWA 430
DB 431 VYAEBAKGMTHHMKRMSPSPMA 454

RESULT 8
BPHA_COMTE STANDARD: PRT: 457 AA.
ID BPHA_COMTE
AC 046372;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
dioxigenase).
CN BPHA.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
NCBI_TaxID=285;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=B-356;
RA MEDLINE=97045812; PubMed=8890734;
RA Sylvestre M., Stirois M., Hurtubise Y., Bergeron J., Ahmed D.,
```

```
RA Shareck F., Barriault D., Guillemette I., Juteau J.M.;
RT "Sequencing of Comamonas testosteroni strain B-356-
RT biphenyl/chlorobiphenyl dioxygenase genes: evolutionary relationships
RT among Gram-negative bacterial biphenyl dioxygenases."
RL Gene 174:195-202(1996).
CC -1 CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1 COPACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -1 PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1 SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase
CC (Bphg) must be present to obtain activity (By similarity).
CC -1 SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U47637; AAC44526.1; -.
DR HSPD: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KM Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxygenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 457 AA; 51691 MW; D133FC0635FACBF5 CnC64;

Query Match 25.5%; Score 615; DB 1; Length 457;
Best Local Similarity 33.1%; Pred. No. 7.2e-41;
Matches 144; Conservative 73; Mismatches 172; Indels 46; Gaps 13;

QY 20 IHGDELFQHELKTIIPARNWMLFTHDSLIPAPGDYVYAKMGIDEVIVSRONGSIRAFIN 79
DB 39 IYADQDLVQLELEVERFGSRWMLLGHETHIPKIGDYLTYMGEDPYIYVRQKQSLKVFNL 98
QY 80 VCRHNGKTLVSEAGNAKGFVCSYHGKGFSGNGELQSVPEKDLGESLNKCLGLEKVE- 138
DB 99 QCRHGRMIRIVSRDGNNAKAFCTYHGMAVYDIAGNLVNVPEKEARCDKKEGDC-GEFKAD 157
QY 139 -----ARVSEFHGFIYGCEDQAPRLMDYLGDAAWYLEPMF-KHSGGLEIVGPPGVYIK 192
DB 158 WGPLDARVETTKGLVFANMDPEAPRLKTYLSDAMPYIMVMDLRTAGTGALIGIOKWIIP 217
QY 193 ANWKAPEWVGDAYHVG-WTHASSLRSGESIFSSLAGNAALPPEGAGLQMT-----SKYG 247
DB 218 CNMKFAAQFCSDMYHAGTTHLSGLAG-----LPPE--MDLTQQLSKNG 262
QY 248 SGKGVLMNGY-----SGVHSADLVPELNAF--CGAKQERLNKEIDVR-ARIYRSHL 296
DB 263 NQFRSAGWGHAGWINDSILLSVGPKITQYWGGAARAKAARVQPLDPLDFGQHM 322
QY 297 NCTVFPNNNSMLTSGVFPVWNPIDANTEVWYTAIVERKMPEDLKRRLADSVOPTTGPAG 356
DB 323 --TYEPTCSFLPGITITITWHPRGNEVEVNAFVYLVADADAEEDIEERRLRNIRKFFNAG 380
QY 357 FWESDDNDNMETASONGKYSRDSLLSLNGFGEDYVGDVAVPGVVGKSA--IGETSYR 414
```

```

Db 381 VEEODDENVEIORVGRHAKSTSLCAKGLVNPKNNDPAY--GKTAYVYAEAR 437
OY 415 GFYRAYOAHVSSNM 429
| : : : :
Db 438 GMYHHMSRMSEPSN 452

RESULT 9
BPHA_PSEPS
ID BPHA_PSEPS STANDARD: PRT: 458 AA.
AC 052028:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
dioxygenase).
GN BPHA OR BPHAL
OS Pseudomonas pseudocataligenes.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=330:
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=KF707;
RC MEDLINE=92165849; PubMed=1537863;
RA Taira K., Hirose J., Hayashida S., Furukawa K.;
RT Analysis of bph operon from the polychlorinated biphenyl-degrading
RT strain of Pseudomonas pseudocataligenes KF707."
RL J. Biol. Chem. 267:4844-4853(1992).
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (BphF) and a ferredoxin reductase
CC (BphG) must be present to obtain activity (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M63673; AAA25743.1; -.
CC HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51437 MW; A6123AD09F97E482 CRC64;

Query Match 25.5%; Score 615; DB 1; Length 458;
Best Local Similarity 33.1%; Pred. No. 7.2e-41;
Matches 147; Conservative 70; Mismatches 183; Indels 44; Gaps 13;

OY 11 ESLSOKHLHGDELFQHELKTLTFARNMGLFVTHDSLIPAPGCVYVTKMGIDEVYVRON 70
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 31 EKGLLDPR-IVADOSLYELELERYFGSRWLLGHESHVPETGDFLATYMGEDPVPVVRK 89
OY 71 DGSIRAFNLNCRHNGKTLVSEAGNARGFVCSYHMGFGSNGELOSVPFKDLGYESLAK 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 90 DKSITKVFNLNCRHNGKTLVSEAGNARGFVCSYHMGFGSNGELOSVPFKDLGYESLAK 149
OY 131 KCLGLKEV-----ARVSEFHGFIYCGFDQEPAPLMYLDGDAAMYLEPMKHS-GGLETV 183
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 DC-GFDAEMGVPLOARVATYKGLVFANMVDYQAPDLETYYLDARPYMDVMDRPAQTVAI 208
OY 184 GPPKGVYIKANWKARPAENFGDAIHVG-WTHASSLRCGESIFPSLAGNALPREG--AGL 240
Db 209 GGMOKWVIPCWKPPAAQFCSDMYHAGTMSHLSGIIAG-----MPREMDLSHA 256
OY 241 QMTSKYSGMGVLWDGYSGVHSDVPELMAFGAK-----QERLNKEIGD 286
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 QVPTK-GNPFAGWGSGHGFVDEPGMLAVNCPKTYQYWTGEPADLAEGQLGITH-P 314
OY 287 VRARIYSHLNCITVFPNNSMLTSGVFKVNPIDANTTEVYTAIYEKMPEDLKRRLAD 346
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 VR-RMGQHM--SVFPTCSFLPAINTIRTWHPROPNEIYVMAFTLVADAPAEIKREYRR 371
OY 347 SVQRTGPAQFWESDDNMETASQNGKTKYQSRSDLSNLGEGEDVYGDVYVPGVYGS 406
Db 372 HNITFSAGVGFEODDENVEIOLGRLGYKAKSQPLNAQMGIGRSOTGHPDPGAVG-Y 430
OY 407 AIGETSYRGFYRAYOAHVSSNMA 430
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 431 VYAEFAARGMTHHMSRMSEPSNA 454

RESULT 10
BNZA_PSEPU
ID BNZA_PSEPU STANDARD: PRT: 448 AA.
AC P08084:
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3) (P1 subunit).
GN BNZA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303:
RN RP SEQUENCE FROM N.A.
RC STRAIN=BE-81;
RC MEDLINE=86032840; PubMed=3667527;
RA Title S., Dol S., Yorifuji T., Takagi M., Yano K.;
RT Nucleotide sequencing and characterization of the genes encoding
RT benzene oxidation enzymes of Pseudomonas putida."
RL J. Bacteriol. 169:5174-5179(1987).
CC -1- CATALYTIC ACTIVITY: Benzene + NADH + O(2) = cis-1,2-
CC dihydrobenzene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BNZA AND BNZB), A
CC FERREDOXIN (BNZC) AND A FERREDOXIN REDUCTASE (BNZD).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M17904; AAA25735.1; -.
CC PIR: A29830; A29830.
DR HSSP: P23094; INDO.

```

```

DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIPOXASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxxygenase; NAD.
FT METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 222 IRON (BY SIMILARITY).
FT METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 448 AA; 50208 MW; 6D1DBD5B2F040BET CRC64;

Query Match 24.6%; Score 593.5; DB 1; Length 448;
Best Local Similarity 33.2%; Pred. No. 3,4e-39;
Matches 143; Conservative 68; Mismatches 175; Indels 45; Gaps 11;

QY 20 IHGDELFQHELTIPARWMLFLTHDSLIPAPGDYVTAKEGIDEVTSRONDSITAFIN 79
DB 35 IYTDDELQLELEFVARSMLLGHETQIRKPEDYITTTMGEDPVYVVRQKASIAVFLN 94
QY 80 VCRHNRKTLVSVEAGNAKGFVCSYHMGFGSNGELQSVPEFKDLGSLNKKCLGKE-- 137
DB 95 QCRHNRMRICRSDAGNAKAFCTSYHGMAVDTAAGNLVNPYEKAESFA-----CLNKKEMS 148
QY 138 --VARVESHGFIYGCDEAPPLMDYLGDAWYLEPMF-KHSGGLELVGPCKVYIKAN 194
DB 149 PLKAREYTKGLIFANWMDENAVDLDTYLGAEAFYMDMLDRFEAGTEALPGVQKWVIPCN 208
QY 195 WKPAENFVGDAYVHVGM--HASSLSRSG--ESIFSSLAGNAALPPEAGLOMTSKYSGMG 251
DB 209 WKPAEQFCSDMYHAGTSHLSGLAGLEPDL-----EKADLAPRVYQYRASMGCHS 263
QY 252 VLMGDSYGVSHADLYPELMAF--GGAKOERLKEIGDVR--ARIYRSHLNCVFPNNM 307
DB 264 GFVYGDPNMLAIMGPKVTSYVTEGPASEKAAERLGSVERGSKLMEWH--TVFPTCSPL 321
QY 308 TCSGVKVPNPIDANTTEVMTAIVAKDMPEDLKRRLADSVQ-----RTTGAGFW 358
DB 322 PGINTVKTLASRAERGEVWAFVVDADAPDDIKEFRARLRFPSVACSSRTGTG-- 379
QY 359 ESDNDNMETASONGKKYOSRSDLSNLGFGEDVYGDVYPGVYKSAIGETSYGFR 418
DB 380 -----SRSTSCSEARSRSRPFNAEMSDQYDNDVYIGRISNNYSEBAAGLIA 430
QY 419 AVQAHVSSSNW 429
DB 431 HMLRMWTSPPDW 441

RESULT 11
BPA1_PSESI STANDARD; PRT: 458 AA.
AC 052438;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxxygenase alpha subunit (EC 1.14.12.18) (biphenyl 2,3-
DE dioxxygenase).
GN BPAH1.
OS Pseudomonas sp. (strain KKS102).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=307;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94324977; PubMed=8048958;
RA Fukuda M., Yasukouchi Y., Kikuchi Y., Nagata Y., Kimbara K.,
RA Horuchi H., Takagi M., Yano K.;
RT "Identification of the bphA and bphB genes of Pseudomonas sp. strains
RT KKS102 involved in degradation of biphenyl and polychlorinated

```

```

RT biphenyls.";
RL Biochem. Biophys. Res. Commun. 202:850-856(1994).
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- CORACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1- SUBUNIT: Heterodimer consisting of 3 BphA1 subunits and 3
CC BphA2 subunits. A ferredoxin (BphA3) and a ferredoxin reductase
CC (BphA4) must be present to obtain activity (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D17319; BA004137.1; -.
DR HSPSP; P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIPOXASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxxygenase; NAD.
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 103 103 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 121 121 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 124 124 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 234 234 IRON (BY SIMILARITY).
FT METAL 240 240 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51876 MW; 128E8C38E2A3CEFAA CRC64;

Query Match 24.3%; Score 585; DB 1; Length 458;
Best Local Similarity 33.6%; Pred. No. 1.6e-38;
Matches 147; Conservative 65; Mismatches 174; Indels 52; Gaps 14;

QY 20 IHGDELFQHELTIPARWMLFLTHDSLIPAPGDYVTAKEGIDEVTSRONDSITAFIN 79
DB 40 IYADDDLEIELERIFARSWLLLGHFAHLPKTDYITTTMGEDPVYVVRQKDGSIKVFILN 99
QY 80 VCRHNRKTLVSVEAGNAKGFVCSYHMGFGSNGELQSVPEFKDLGSLNKKCLGKEV- 138
DB 100 QCRHNRMRICRSDAGNAKAFCTSYHGMAVDTAAGNLVNPYEKAEPDCKKEGDC-GPDKAD 158
QY 139 -----ARVSEHGFYGCDEAPPLMDYLGDAWYLEPMF-KHSGGLELVGPCKVYIK 192
DB 159 WGPLQARVETYKGLIFANWDAEAPDLKTYLSDAMPYMDMLDRTEAGTVYVGMQKWTIP 218
QY 193 ANMKPAENFVGDAYVHV--WTHASSLSRSGESIFSSLAGNAALPPEAGLOMT-----SKY 247
DB 219 CWKRAAEQFCSDMYHAGTSHLSGLAGLEPDL-----SLPPE-----MDLVQVMSKNG 263
QY 248 SGMGYLMDYSGVSHADLYPELMAFGAK-----QERLNRKEIGDVRAR-IYRSHL 296
DB 264 SQFRAMGSHGSGWFLINDAAILMAVWPKITQYWGQPAKAAKARLNRQMPQTMGQGM 323
QY 297 NCTVFPNNMGLTCSGVFKWNPIDANTTEV---WTYAIYEKMPEDLKRRLADSVQRTTG 353
DB 324 --TVFPTCSFLPGINTIRSWHPRGNEVECGPSWSMPAR---PDIKEFRRONIRTFN 378
QY 354 PAGFESDNDNMETASONGKKYOSRSDLSNLGFGEDVYGDVYPGVYKSA--IGET 411
DB 379 AGCTEPDODDGEWVLEQRLGHKAKSAFLCAOMGLNVPNKSNDPFP---GKTAIYVABE 435
QY 412 SYRGFYRAYQAHVSSSNW 429

```

Db 436 AARGMYHHWMMSEPNW 453

# RESULT 12

XYLX\_PSEPU STANDARD; PRT; 454 AA.

ID XYLX\_PSEPU

AC P23099;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Toluene 1,2-dioxygenase alpha subunit (EC 1.14.12.-).

GN XYLX.

OS Pseudomonas putida.

OC Plasmid TOL PMW0.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas

OX NCBI\_TaxID=303;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92155191; Pubmed=1740120;

RA Neidle E.L., Hartnett C., Ornston L.N., Bairoch A., Reikik M.,

RT C15-diol dehydrogenases encoded by the TOL PMW0 plasmid xylX gene

and the Acinetobacter calcoaceticus chromosomal bend gene are members

of the short-chain alcohol dehydrogenase superfamily."

RL Eur. J. Biochem. 204:113-120(1992).

CC -1 COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.

CC -1 SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS. THE

ELECTRON TRANSFER COMPONENT (XYLX), AND AN

-1 SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M64747; AAA26047.1; -

DR PIR: S23482; S23482.

DR InterPro: IPR001281; Rieske.

DR InterPro: IPR001663; Ring\_hydroxyl\_A.

DR Pfam: PF00848; Ring\_hydroxyl\_A.

DR PRINTS: PR00090; RINGDIOXGNASE.

DR PROSITE: PS00570; RING-HYDROXYL ALPHA: 1.

KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;

KM Dioxigenase; NAD; Plasmid.

FT METAL 92 92 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT METAL 94 94 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT METAL 221 221 IRON (BY SIMILARITY).

FT METAL 226 226 IRON (BY SIMILARITY).

SO SEQUENCE 454 AA; 51898 MW; F68D1C41E236C077 CRC64;

Query Match 17.0%; Score 410.5; DB 1; Length 454;

Best Local Similarity 31.2%; Pred. No. 8.3e-25;

Matches 105; Conservative 64; Mismatches 135; Indels 33; Gaps 14;

QY 23 DEELFOHELTIFARNKLEFLTHDSLIPAGDGYVAKMGIDEIVSRONDGIRAFANCR 82

Db 34 DPLRLFLEHMKIEGNNIYLAHESQIPEKNDYTTOMGRQPIFTIRNKGELNFAVNACS 93

QY 83 HRKGLTVSAGNAKAGVCSHGSGNSGELQSVPEKDL-YGESLKKKLG---LKEY 138

Db 94 HRGATLCRFRSGNAKATHTCSFHGWTFSNSGKLLKVKDKPGAGYDPSFD--CDGSHDLKKV 151

QY 139 ARVESFHGFIYCEPDQEAAPLMDYLGDAWYLEBPMFKHS-GGLELYGPPCKVYIKANMKA 197

Db 152 ARFASRYGFLFGLSREDVAPLEEFLEGSRKYIDWVVDQSEPLEVNGSTYYEGMKKY 211

QY 198 PAENFVGDAYHV---WTHASS-----LR-SGESIFSSLAGNNAALPPREGLOMTSKYGS 248

Db 212 QVEN-GADGYHVSIVHWNYATQOORKLQDAGDDIRAMTA--SSWGDDGGGFY---SFEK 265

QY 249 GMGVLMDSGVSHSADYLPPELMAFGAKQKORLENGIDVBAR-IYRSHLNCTVFPNNSMU 307

Db 266 GHQWVMAWAGCPKRPPLF-----AERDLASEFGEARADWIGVSRMLCLYPLNYLMA 317

QY 308 TCSCG-VFKVWNPIDANTTEWVTAIVER-DAPEDLKR 342

Db 318 DQFGSQLRITRPLSDVDRREITTYCIAPKGETPRRAR 354

# RESULT 13

BENA\_ACICA

ID BENA\_ACICA

AC P07769;

DT 01-NOV-1991 (Rel. 20, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10).

GN BENA.

OS Acinetobacter calcoaceticus.

OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;

OC Acinetobacter.

OX NCBI\_TaxID=471;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BD413 / ADP1;

RX MEDLINE=91356314; Pubmed=1885518;

RA Neidle E.L., Hartnett C., Ornston N.L., Bairoch A., Reikik M.,

RT "Nucleotide sequences of the Acinetobacter calcoaceticus benABC genes

for benzoate 1,2-dioxygenase reveal evolutionary relationships among

multicomponent oxygenases."

RL J. Bacteriol. 173:5385-5395(1991).

CC (2)

RP REVISIONS TO 84, 103-104, 171-172 AND 380-382.

RA Elby D.M., Neidle E.L.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

CC -1 FUNCTION: DEGRADATION OF BENZOATE TO 2-HYDRO-1,2-DIHYDROXYBENZOATE

(DHB).

CC -1 CATALYTIC ACTIVITY: Benzoate + NADH + O(2) = catechol + CO(2) +

NAD(+).

CC -1 COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.

CC -1 PATHWAY: DEGRADATION PATHWAY LEADING FROM BENZOATE TO CATECHOL.

CC -1 SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS: THE

TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BENA AND BENB), AND

AN ELECTRON TRANSFER COMPONENT (BENC).

CC -1 SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AF009224; AAC46436.2; -

DR PIR: S23477; S23477.

DR InterPro: IPR001281; Rieske.

DR InterPro: IPR001663; Ring\_hydroxyl\_A.

DR Pfam: PF00848; Ring\_hydroxyl\_A.

DR PRINTS: PR00090; RINGDIOXGNASE.

DR PROSITE: PS00570; RING-HYDROXYL ALPHA: 1.

KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;



```

KW Dioxigenase: NAD.
FT METAL 95 95 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 97 97 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 118 118 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 224 224 IRON (BY SIMILARITY).
FT METAL 229 229 IRON (BY SIMILARITY).
SQ SEQUENCE 461 AA: 52228 MW: 52228 MW: CFC3247A3C4379 CRC64:

Query Match 16.5%; Score 397; DB 1; Length 461;
Best Local Similarity 27.7%; Pred. No. 9.7e-24;
Matches 128; Conservative 78; Mismatches 204; Indels 52; Gaps 20;

OY 5 NRIIV--SESGSOKH-LIHGDEELFOHELTIFARNWLFTHDSLIPAPGYVAKMGI 61
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 16 DELVDNTETGEFKLRSTFQDALEFMKTIIEGNNWYLLHESQIPNNNDYITTYIER 75
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 62 DEVIVSRQNDGSIIRAFNLVCRHGRKTVLSEVAGNAKGFVCSYHGFGSGELQSVPEEK 121
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 76 OPILLARNNGELNMINACSHRGAQLCRHKGKNTTYTCPFHGWTFNNSKLTKVDPIS 135
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 122 DL-YGSLKKK-CLGJKEVARVESFHGFIYCGFQDEAPRLMYLGDAMYLEPMKHSK- 178
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 136 DAGYSPCEVDGSHDLKRYARESYKGFLEGLNDPVSLOEFLGETTKIIMYGOQSDQ 195
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 179 GLEIVPPEKVVIVKANMKAPENFVGDAYHVG---WTHASSL-----RSGESIFSLAG 229
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 196 GLEVLGVSTYTYEGNMKLTALEN-GADGYHVSANHNRYATIQHREKQAGOTIRAMSG 254
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 230 NAALPEPAGLQMTSKYK--SGMGVLMDSYSGVSHASDLVPELM---AFGAKOERLNKE 283
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 255 SMG--KHGGG-----SYGREGHMLMTQGNPDEPNPKAAYETEKGAAMSKWM--- 304
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 284 IDGVRRIRIRSHLNTVFPNNSMLTCSG-VKVMNPIDANTTEWYTAIVE-KDMPEDLK 341
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 305 -----IERSR-NCLXPNVYLMQFGSOLVLRISVKNKEVTYICAPVGEAPEARA 356
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 342 RRLDSVORTTGPAGFWEEDNDNDNETASONGKTYOSRSDLSLNGFEGDYGG--DA- 397
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 357 RRIR-QYEDFPMASGMATPDDLEFRACQAGYAGIELENDNCR--GSRHWITGPDADA 413
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 398 ---VYPGVKSAIGETSYRGFRAYVAHVSSNMAFEFHAS 436
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 414 EIGLKPALISGITEDEGLYLAOHYWLKSMKQAIAAKEEFAS 455
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14
YEAH_ECOLI STANDARD: PRT: 374 AA.
AC P76253;
DT 15-JUL-1998 (Rel. 36, created)
DT 15-JUL-1998 (Rel. 36, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Putative dioxigenase alpha subunit yeast (EC 1.14.1.1).
GN YEAW OR B1802 OR Z2845 OR ECS2511.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OC NCBI_TaxID=562, 83334;
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete sequence of Escherichia coli K-12." ;
RT Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;

```

```

RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Maki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakada S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasubdaram S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map." ;
RT DNA Res. 3:379-392(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocitis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." ;
RT Nature 409:529-533(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsuno E., Nakayama K., Murota T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa G., Ogasawara N., Yasunaga T.,
RA Kudara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12." ;
RL DNA Res. 8:11-22(2001).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- SUBUNIT: PROBABLE HETERODIMER OF YEAW AND YEAX.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC
CC EMBL: AE000274; AAC74872.1; -
CC EMBL: D90823; BAA15597.1; ALT_INIT.
CC EMBL: D90824; BAA15606.1; ALT_INIT.
CC EMBL: AE005403; AAC56791.1; -
CC EMBL: AP002558; BAB35934.1; -
CC Ecogene: EG13509; yeaw.
CC InterPro: IPR001281; Rleske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rleske; 1.
CC PROSITE: PS00570; RING-HYDROXYL ALPHA: FALSE NEG.
KW Hypothetical protein; Oxidoreductase; Iron-sulfur; Iron; Dioxigenase;
KW NAD; Complete proteome.
FT METAL 89 89 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 91 91 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 109 109 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 211 211 IRON (BY SIMILARITY).
FT METAL 216 216 IRON (BY SIMILARITY).
SQ SEQUENCE 374 AA: 42561 MW: BB5386ACA9585606 CRC64;

Query Match 10.2%; Score 246; DB 1; Length 374;
Best Local Similarity 26.6%; Pred. No. 5.3e-12;
Matches 98; Conservative 53; Mismatches 127; Indels 90; Gaps 19;

```

```

OY 23 DEELFOHELTIFARNWLFTHDSLIPAPGYVAKMGIDEVIYSRQNDGSIIRAFNLVCR 82
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 31 DONAFHEKENVFAKSWICVASHSELANANDYVTRIEIGESIVLVRGDKVLAIFYNVC 90
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:55:37 ; Search time 13.1667 seconds  
(without alignments)  
3278.305 Million cell updates/sec

Title: US-09-843-250-34

Perfect score: 2409

Sequence: 1 MNVNNKILVSESGLSQKHLI.....AEFFHASSTWHTLTKTDR 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2402	99.7	449	2 JS0071	naphthalene dioxyg
2	2402	99.7	449	2 JS0071	naphthalene 1,2-di
3	2344	97.3	449	2 JUN0644	naphthalene 1,2-di
4	2280	95.1	449	2 C55217	polycyclic aromati
5	2051	85.1	447	2 T31134	2-nitrotoluene dio
6	955	39.6	459	2 T31134	naphthalene dioxyg
7	772	32.0	450	2 T31256	terminal oxygenase
8	735	30.5	453	2 A65031	biphenyl dioxygena
9	735	30.5	453	2 D91054	biphenyl dioxygena
10	735	30.5	453	2 G85898	biphenyl dioxygena
11	723.5	27.0	450	2 T31258	aromatic oxygenase
12	650.5	27.0	450	1 JN0812	benzene 1,2-dioxyg
13	648.5	26.9	450	1 A36516	toluene dioxygena
14	637	26.4	461	2 S51757	biphenyl dioxygena
15	615.5	25.6	459	1 B41858	biphenyl dioxygena
16	615	25.5	457	1 JCA493	biphenyl dioxygena
17	615	25.5	457	1 A42409	biphenyl dioxygena
18	603	25.0	431	2 JWO098	carbazole dioxygen
19	593.5	24.6	448	1 A29830	benzene 1,2-dioxyg
20	585	24.3	458	2 JCA467	biphenyl dioxygena
21	450	18.7	469	2 T50934	dioxygenase D1AL,
22	422	17.5	471	2 T31281	benzoate 1,2-dioxy
23	410.5	17.0	454	2 A41659	toluene 1,2-dioxyg
24	407.5	16.9	455	2 E83332	anthranilate dioxy
25	392	16.3	461	2 S23477	probable ring-hydr
26	387.5	16.1	464	2 G83331	biphenyl dioxygena
27	336.5	14.0	424	2 E83384	aromatic oxygenase
28	330	13.7	426	2 T31278	
29	312.5	13.0	391	2 T31251	

30	311.5	12.9	468	2 G97447	hypothetical prote
31	311.5	12.9	468	2 A82665	ring hydroxylating
32	291	12.1	429	2 F82970	probable ring hydr
33	287	11.9	412	2 H95311	probable aromatic-
34	283	11.7	415	2 AG3320	benzoate 1,2-dioxy
35	269.5	11.2	404	2 H87635	Rieske 2Fe-2S faml
36	263	10.9	420	2 T31285	biphenyl dioxygena
37	249.5	10.4	374	2 AF0304	probable dioxygena
38	246	10.2	374	2 C85791	probable choline m
39	246	10.2	374	2 G90942	probable choline m
40	246	10.2	374	2 B64941	probable choline m
41	233.5	9.7	382	2 G70946	probable choline m
42	224	9.3	446	2 T14542	probable dioxygena
43	223	9.3	439	2 T09214	choline monooxygen
44	214.5	8.9	2	JC5354	2-nitrotoluene dio
45	186	7.7	426	2 T08550	choline monooxygen

#### ALIGNMENTS

##### RESULT 1

JS0071 naphthalene dioxygenase (EC 1.14.12.-) ndob protein - Pseudomonas putida

C:Species: Pseudomonas putida

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 20-Aug-1999

C:Accession: JS0071

R:Kurkela, S.; Lehtvaestaho, H.; Palva, E.T.; Teeri, T.H.

Gene 73, 355-362, 1988

A:Title: Cloning, nucleotide sequence and characterization of genes encoding naphthal

A:Reference number: JS0070; MUID:89211973; PMID:3243438

A:Accession: JS0071

A:Molecule type: DNA

A:Residues: 1-449 <KUR>

A:Cross-References: GB:M23914; NID:9151392; PIDN:AAB47591.1; PID:9151394

C:Comment: Naphthalene dioxygenase system is composed of three proteins.

C:Genetics:

A:Gene: ndob

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [

C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein

F:71-119/Domain: Rieske [2Fe-2S] homology <RS>

F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 99.7% Score 2402; DB 2; Length 449;

Best Local Similarity 99.8%; Pred. No. 4.8e-183.

Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MNVNNKILVSESGLSQKHLIHGDELFQHELTIFARNMLFTHDSLIPAGDYVTAKMG	60
DB	1	MNVNNKILVSESGLSQKHLIHGDELFQHELTIFARNMLFTHDSLIPAGDYVTAKMG	60
QY	61	IDEVYVSRONDSIRAFVWNCRRKGTIVSVENAKKGVCSYHGWFSGNGLSVPE	120
DB	61	IDEVYVSRONDSIRAFVWNCRRKGTIVSVENAKKGVCSYHGWFSGNGLSVPE	120
QY	121	KDLVGSLEKKGICGLKEVARVSEFHGFIYCFDOEAPLMDVLGDAAWLEPFKHSGL	180
DB	121	KDLVGSLEKKGICGLKEVARVSEFHGFIYCFDOEAPLMDVLGDAAWLEPFKHSGL	180
QY	181	ELVGPGRKVIKANKKAPAEFVGDAYHVGWTHASSLSRGSFISSLAGMAALPREGAL	240
DB	181	ELVGPGRKVIKANKKAPAEFVGDAYHVGWTHASSLSRGSFISSLAGMAALPREGAL	240
QY	241	QMTSKYSGMGVLMGYSVSHSADLVPELMAFGAKQELNKEIGDYRKRIYRSHLNCIV	300
DB	241	QMTSKYSGMGVLMGYSVSHSADLVPELMAFGAKQELNKEIGDYRKRIYRSHLNCIV	300
QY	301	PNNNSMLTSCGVFKWNPIDANTTEWYTAIVEKMDPELKRRLDSVORTGPAFGFWS	360
DB	301	PNNNSMLTSCGVFKWNPIDANTTEWYTAIVEKMDPELKRRLDSVORTGPAFGFWS	360
QY	361	DDNDNMETAQNGKKYQSHSDLSNLGFGEDVYGDVAVPGVVGSAIGETSYRGFFYRA	420
DB	361	DDNDNMETAQNGKKYQSHSDLSNLGFGEDVYGDVAVPGVVGSAIGETSYRGFFYRA	420

Db 361 DDNDMETSANGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFERHASTTWHTELTKTTDR 449  
Db 421 QAHVSSSNMAEFERHASTTWHTELTKTTDR 449

## RESULT 2

S27632

naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large chain  
C:Species: Pseudomonas sp.  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 15-Oct-1999  
C:Accession: S27632; B49343

R:Denome, S.A.; Young, K.D.  
Submitted to the EMBL data library, February 1992  
A:Description: Cloning and molecular characterization of genes involved in metabolism of  
A:Reference number: S27631  
A:Accession: S27632

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <DEN>  
A:Cross-references: EMBL:M60405

R:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
J. Bacteriol. 175, 6890-6901, 1993  
A:Title: Metabolism of dibenzothiophene and naphthalene in Pseudomonas strains: complete  
A:Reference number: A49343; MUID:94042852; PMID:8226631  
A:Accession: B49343

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 14-462 <DE2>  
A:Cross-references: GB:M60405; NID:g151195; PIDN:AAA16125.1; PID:g294351  
A:Experimental source: strain C18  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:84-132/Domain: Rieske [2Fe-2S] homology <RSK>  
F:94,96,114,117/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 99.7%: Score 2402; DB 2; Length 462;  
Best Local Similarity 99.8%; Pred. No. 5e-183;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNKKLIVSSGSLQKHLHGDEELFOHELKTIFARNWLELTHDSLIPAPGDIYTAAMG 60  
Db 14 MNYNKKLIVSSGSLQKHLHGDEELFOHELKTIFARNWLELTHDSLIPAPGDIYTAAMG 73  
QY 61 IDEVTVSRONGSGIRAFNLNCRHKGKTLVSEAGNAKGFVCSYHGWFSGNGELQSVPE 120  
Db 74 IDEVTVSRONGSGIRAFNLNCRHKGKTLVSEAGNAKGFVCSYHGWFSGNGELQSVPE 133  
QY 121 KDLYGESLNKKCLGKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 134 KDLYGESLNKKCLGKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 193  
QY 181 ELYGPPGKVVYIKAMWKAENFVGDAHVGTWTHASSLSRSGSIFSSLAGNAALPREGAGL 240  
Db 194 ELYGPPGKVVYIKAMWKAENFVGDAHVGTWTHASSLSRSGSIFSSLAGNAALPREGAGL 253  
QY 241 QMTSKYSGMGVLMDDGYSGVNSADLVPPELMAFGAKOERLKEIGDVARARYSHLNCVY 300  
Db 254 QMTSKYSGMGVLMDDGYSGVNSADLVPPELMAFGAKOERLKEIGDVARARYSHLNCVY 313  
QY 301 FPNNSMLTCSGVFKWNPIDANTTEVMTYALVEKDMPEDLKRLADSVORTTGPAGFWE 360  
Db 314 FPNNSMLTCSGVFKWNPIDANTTEVMTYALVEKDMPEDLKRLADSVORTTGPAGFWE 373  
QY 361 DDNDMETSANGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 374 DDNDMETSANGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 433  
QY 421 QAHVSSSNMAEFERHASTTWHTELTKTTDR 449  
Db 434 QAHVSSSNMAEFERHASTTWHTELTKTTDR 462

## RESULT 3

JN0644

naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large cha  
N:Alternate names: nabc protein  
C:Species: Pseudomonas putida  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
C:Accession: JN0644

R:Simon, M.J.; Osslund, T.D.; Saunders, R.; Ensley, B.D.; Suggs, S.; Harcourt, A.; Su  
Gene 127, 31-37, 1993  
A:Title: Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida st  
A:Reference number: JN0640; MUID:93252277; PMID:8486285  
A:Accession: JN0644

A:Molecule type: DNA  
A:Residues: 1-449 <STM>  
A:Cross-references: GB:M83949; NID:g151384; PIDN:AAA25902.1; PID:g151387  
C:Comment: This protein is a member of naphthalene dioxygenase multicomponent enzyme  
C:Genetics:

A:Gene: nabc  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2Fe  
C:Keywords: 2Fe-2S; heterotetramer; metalloprotein; oxidoreductase; Rieske iron-sulfu  
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 97.3%: Score 2344; DB 2; Length 449;  
Best Local Similarity 96.7%; Pred. No. 2e-178;  
Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 MNYNKKLIVSSGSLQKHLHGDEELFOHELKTIFARNWLELTHDSLIPAPGDIYTAAMG 60  
Db 1 MNYNKKLIVSSGSLQKHLHGDEELFOHELKTIFARNWLELTHDSLIPAPGDIYTAAMG 60

QY 61 IDEVTVSRONGSGIRAFNLNCRHKGKTLVSEAGNAKGFVCSYHGWFSGNGELQSVPE 120  
Db 61 IDEVTVSRONGSGIRAFNLNCRHKGKTLVSEAGNAKGFVCSYHGWFSGNGELQSVPE 120

QY 121 KDLYGESLNKKCLGKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 121 KDLYGESLNKKCLGKEVARVESFHGFIYGCDFDEAPPLMDYLGDAAWYLEPMFKHSGGL 180

QY 181 ELYGPPGKVVYIKAMWKAENFVGDAHVGTWTHASSLSRSGSIFSSLAGNAALPREGAGL 240  
Db 181 ELYGPPGKVVYIKAMWKAENFVGDAHVGTWTHASSLSRSGSIFSSLAGNAALPREGAGL 240

QY 241 QMTSKYSGMGVLMDDGYSGVNSADLVPPELMAFGAKOERLKEIGDVARARYSHLNCVY 300  
Db 241 QMTSKYSGMGVLMDDGYSGVNSADLVPPELMAFGAKOERLKEIGDVARARYSHLNCVY 300

QY 301 FPNNSMLTCSGVFKWNPIDANTTEVMTYALVEKDMPEDLKRLADSVORTTGPAGFWE 360  
Db 301 FPNNSMLTCSGVFKWNPIDANTTEVMTYALVEKDMPEDLKRLADSVORTTGPAGFWE 360

QY 361 DDNDMETSANGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 361 DDNDMETSANGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420

QY 421 QAHVSSSNMAEFERHASTTWHTELTKTTDR 449  
Db 421 QAHVSSSNMAEFERHASTTWHTELTKTTDR 449

RESULT 4  
C55217  
polycyclic aromatic hydrocarbon dioxygenase (EC 1.14.12.-) iron-sulfur protein large

C:Species: Pseudomonas putida  
C:Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 20-Jun-2000  
C:Accession: C55217

R:Takizawa, N.; Kaide, N.; Torioge, S.; Moritani, T.; Sawada, T.; Satch, S.; Kiyohara  
J. Bacteriol. 176, 2444-2449, 1994  
A:Title: Identification and characterization of genes encoding polycyclic aromatic by

A:Reference number: A55217; MUID:94209249; PMID:8157615  
A:Accession: C55217  
A:Status: preliminary



Db 186 LKSTLACNMKVPTEFNEVGDYHGWTHAALQIMIGGELAGLSGNRPADPFDDLGLOFTMR 245  
QY 246 YGSGMGVLMIDYSGVH--SADLYPELMAFGAKOERLNKEIGDVARIRYRSHLCTVPPN 303  
Db 246 HGHGFGIIDAATAIHKRCGYVKKYLEETRGCIKEKCP-----REKLYGHMWTSLIPN 301  
QY 304 NSMLTCSGVKVPNPIDANTTEVWTVAIVEKMDPEDLKRLADSVORTTGPAGFWESODN 363  
Db 302 CSFLYGNTFTFKIMHPRGPHIEVWTYTMVPRKNADTETKRSTIORAIRSFAGTGLESDDG 361  
QY 364 DNNETASQNKKKYQSRDLSNLGFGEDVYG-DAVYPGVYKSAIGETSTRGCFRAQA 422  
Db 362 ENNSATYNNNGNITTRGRM--NSSMKDRGCPHPVYPGIVGSFICETSTRGCFREMOE 419  
QY 423 HVSSSNMAEFEEHASTW 439  
Db 420 MLDAPDMAIRANDDTW 436

## RESULT 7

T31256 terminal oxygenase component large chain homolog - Sphingomonas aromaticivorans plasmid  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31256  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A:Reference number: Z20992  
A:Accession: T31256  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-450 <ROM>  
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378397; PIDN:AD03980.1  
C:Genetics:

A:Genome: plasmid pML1  
A:Note: Dphala  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
F:81.83,101.104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 32.0%; Score 772; DB 2; Length 450;

Best Local Similarity 36.1%; Pred. No. 1.8e-53;

Matches 157; Conservative 88; Mismatches 184; Indels 6; Gaps 6;

QY 1 MNYNNKILVSESGSKQHLHGDDELFOHELKTIFARNMLFLTHDSILPAPGDVYAKMG 60  
Db 1 MNSIDLVDSRGRORSIYASDIYROELERIFGKWLFLVHTSQIPKGDVFRFTMG 60  
QY 61 IDEVIVSRQDGSIRAFNLVCRHRGKTLVSVEAGNAKGFVCSYHGWGFGSNGELQSVPEE 120  
Db 61 EDVIVYRQDGSIRAFNLVCRHRGKTLVSVEAGNAKGFVCSYHGWGFGSNGELQSVPEE 120  
QY 121 KDLVGSLSNKKCGLKEVAVVESEFHGFIYGCFOEAPRLMDYLGDAAWYLEP-KFKHSGG 179  
Db 121 NEVYFEGELDTKFGLLPVTVVAEYKGLVFGCWDANSPLDYLGDAAWYLEP-LDAMPGG 180  
QY 180 LELVGPCKGVKIRANKKAPAEFNEVGDYHGWTHAALQIMIGGELAGLSGNRPADPFDDLGLOFTMR 245  
Db 181 SALLGETQKKNVLTGNKFLPVENVCBGDGYHLGMAHAGMAAQAQSDLTGLSGNSGVLDLG 240  
QY 238 AGLOMTSKYSGMGVLMIDYSGVHSAADLVPELMAFGAKOERLNKEIGDVARIRYRSHL 296  
Db 241 -GLSVAGMGHMYLALDGVSAFYDPPILEYLEANRQTVIDRLGEVGRQVMAQV 299  
QY 297 NCYVFFNNNSMLTCSGVKVPNPIDANTTEVWTVAIVEKMDPEDLKRLADSVORTTGPAG 356  
Db 300 NTIFIFNLDLLPLNMRVYHPKPGQIEQWTMAENDPEAKAIILEOQCLTFJGAG 359  
QY 357 FWESDNDNMNETASQNKKKYQSRDLSNLGFGEDVYGDAVYPGVYKSAIGETSTRGCF 416  
Db 360 LFPNDGDNLTACTEBSRGRMTAOMDYTNMALGRSGKRBS-FPGDIAAGLVSEHNRIF 418

QY 417 YRAYQAHVSSNNMAE 431  
Db 419 YRWQEHMAETMAE 433

## RESULT 8

biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escheri  
A:65031  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: A65031; S49292  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A:Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: A65031

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-453 <BLAT>

A:Cross-references: GB:AE000340; GB:U00096; NID:g1788883; PIDN:AAC75591.1; PID:g17888

A:Experimental source: strain K-12, substrain MG1655

R:Tutlin, E.; Gasser, F.; Biville, F.

A:Description: Cloning and sequencing of an E. coli gene homologous to dioxygenase of

A:Reference number: S49292

A:Accession: S49292

A:Molecule type: DNA

A:Residues: 1-19, 'A', 21-383, 'ATAATATNCVKKW', 397-398, '451', 'R', '453', 'SNAATTAIALTLTFOKPL

A:Cross-references: EMBL:Z37966; NID:g550593; PIDN:CA86018.1; PID:g550596

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske I

C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein

F:75-123/Domain: Rieske [2Fe-2S] homology <RSK>  
F:85,87,105,108/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 30.5%; Score 735; DB 2; Length 453;

Best Local Similarity 36.4%; Pred. No. 1.6e-50;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGSELFQHELKTIFARNMLFLTHDSILPAPGDVYAKMGIDEVIVSRQDGSIRAFIN 79  
Db 24 IYTPDITQLEKRIFGKWLFLVHTSQIPKGDVFRFTMG 83  
QY 80 VCRHRGKTLVSVEAGNAKGFVCSYHGWGFGSNGELQSVPEEKLYGSLNKKCLGKEVA 139  
Db 84 QCRHRARVSYADCGNTRAFPCRYHGWGSGINGELDIDVLEPRAPRGCLCKSHMGLEVP 143  
QY 140 RVESFHFQIYGCFOEAPRLMDYLGDAAWYLEP-KHSGLELVGPCKGVYIKAWKAP 198  
Db 144 CVESYKGLIFGNMDSAPGLNDYLDIAWYLDGMDLRREGTEIVGVQKVTINCMKPEP 203  
QY 199 AENFVGDYHGWTHAALQIMIGGELAGLSGNRPADPFDDLGLOFTMR 245  
Db 204 AEOFASDQYHALFHSASAVOYLKAKKDGSDKRLGD-----GOTARFVETAKKDLQIG 256  
QY 235 --PEGAGLOMTSKYSGMGVLMIDYSGVHSAADLVPELMAFGAKOERLNKEIGDVARIRY 292  
Db 257 ODGSGSGFFEFKEDDANWV--DGAVSSYRETYAE-----AEQRLGEVRLRL 303  
QY 293 RSHLNCYVFPNNSMLTCSGVKVPNPIDANTTEVWTVAIVEKMDPEDLKRLADSVORTT 352  
Db 304 AGHANN--IFPTLSWLNATATLRVWHPRGPOVEWMAFCITDKAASDEVKAFFNSATRAF 361  
QY 353 GPAGFWESDNDNMNETASQNKKKYQSRDLSNLGFGEDVYGDAVYPGVYKSAIGETST 412  
Db 362 GPAGFWESDNDNMNETASQNKKKYQSRDLSNLGFGEDVYGDAVYPGVYKSAIGETST 412  
QY 413 YRAYQAHVSSNNMAEFEEHASTW 439  
Db 421 ARGWYQRMADLSSSEWQEVLDATAAQQDEVK 453

## RESULT 9



```
OY 62 DEVYISRONDSIRAFINVCNRRKTLVSEAGNAKGFVCSYHGWFSGNSELQSVPEK 121
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 69 DAVIVARQPGSIVMLNSCHRNKKNKCFADAGNTRRFVCHVHGAFPTAGDKMHEBY 128
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 122 DLGESLNKKCLGLKEVARVESFHGFIYGCDFQDAPRLMDYLGAAYLLEPMF-KHSGL 180
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 129 CYDAGDDIDFKHNGIKNAKVGNYKGLVFATFNSDAPSLLEAWLGDPRWYLDMLIDNEEGT 188
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 181 ELVGPFGKVVYKAMWKAPEAFVGDVAVHGWTHASSLSRSGSIFSSLAGNALPREGAGL 240
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 189 EFIGGCTKSIVASAMKGFVENFIDGATHAGWTHSGTR-----SMNNGOPFP-----I 237
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 241 QMTRSKYSGM-GVLM-D-GYSGVSHADLV--PELMAFGAKOERLNKEIGDVARIRYRSHL 296
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 238 DMDRSHASVNGHGWEGTEGVGLFLGGRPKVMDYKIRPKMAERLGRKSKIFGSVA 297
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 297 NCIVFPNNKSLTCSGVKRWMPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTTGPAG 356
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 298 SASIFPVSFLPGISTEROMQPKPMQELKTWIVYNNKMPDDIKEEYTKGVQMOTFGPG 357
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 357 FWESDDNDNNETASQNKQKQSDLSNLGFGEDVYGDVYGVGKSAIGETSYRGF 416
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 358 TPEMDDENNENCTTVNRGVYTRHERLHRCGICRQLEHDTL-PEIYRGOYNDANORGF 416
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 417 YRAY 420
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 417 YORW 420
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 12
JN0812
benzene 1,2-dioxygenase (EC 1.14.12.3) terminal oxygenase component large chain - Pseudoc
C:Species: Pseudomonas putida
C>Date: 19-May-1994 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C/Accession: JN0812
R:Tam, H.M.; Tang, H.Y.; Joannou, C.L.; Abdel-Wahab, N.H.; Mason, J.R.
Gene 130, 33-39, 1993
A:Title: The Pseudomonas putida ML2 plasmid-encoded genes for benzene dioxygenase are ut
A:Reference number: JN0810; MUID:93345820; PMID:8344526
A:Accession: JN0812
A:Molecule type: DNA
A:Residues: 1-450 <TAN>
A:Cross-References: GB:L04642; GB:L04643; NID:96552505; PIDN:AAA17758.1; PID:9309855
A:Experimental source: strain ML2
C:Comment: This enzyme is involved in catalyzing the oxidation of benzene to cis-1,2-dih
C:Genetics: The reduced enzyme catalyzes the oxidation of benzene to cis-benzene dihydrod
A:Gene: DedC1
A:Genome: plasmid
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe
C:Keywords: 2Fe-2S; aromatic hydrocarbon catabolism; metalloprotein; oxidoreductase; Rie
F:86-134/Domain: Rieske [2Fe-2S] homology <RSK>
F:96,98,116,119/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 27.0%; Score 650.5; DB 1; Length 450;
Best Local Similarity 34.6%; Pred. No. 8,4e-44;
Matches 148; Conservative 69; Mismatches 174; Indels 37; Gaps 11;

OY 20 IHGDELFQHELTIFARNMLFLTHDSLIPAPGDVYAKMGIDDEVYISRONDSIRAFIN 79
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 35 IYDDEDLQLELEVERFARSWLLGHETHIRKPGDYFTTYMKEDPVVYVVRQDAIAVFLN 94
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 80 VCHRRKTLVSEAGNAKGFVCSYHGWFSGNSELQSVPEKDLGESLNKKCLGLKE-- 137
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 95 QCHRRGRICRSADGAKNAKATTCSHGMAVDTAGNLINVPYAESFA-----CLDKKMS 148
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 138 --VARVESFHGFIYGCDFQDAPRLMDYLGAAYLLEPMF-KHSGLDELVGPFGKVVYKAN 194
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 149 PLKARVETVYKGLIFANNDENALDIDTYLGEAKFYNDHMLDRTEAGTEVYPIQKMWIPCN 208
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 195 WKPAPEFVODAHVGVGT-HASSLSRSGSIFSSLAGNALPREGAGLQMTSKYSGMGVL 253
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
```

```
DB 209 WKFAEDQCSMDYHAGTTAHLGIIAGLPEDLELADLA--P-----KRGQYRAS 257
OY 254 WDGY-SGVHSHAD-----LVPELMAF--GGAKOERLNKEIGDVR--ARIYRSHLCTVFE 301
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 258 WGHSGSGFIYICDPNNMLAMGPKVTSYLTGEPRAEKAERLGSIERGTKMLDEM--TYF 315
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 302 PNNSLTCSGVGVKWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTTGPAGFESD 361
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 316 PTCSEFLPQVNTIRIWHPRGRNEVEVMAFTYVDADAPDDIKEEPRROTLRFFSAGVFEED 375
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 362 DNDNKTASQNGKKYQSDLSNLGFGEDVYGDVYGVGKSAIGETSYGFRATQ 421
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 376 DGEWVEIOLHILRCHAKSRPFNAEMSGQTVNDPDPYGRISNNVYSEEAARGLYAHML 435
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 422 AHVSSSNM 429
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 436 KMTSTPDM 443
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

RESULT 13
A36516
toluene dioxygenase (EC 1.14.12.11) terminal oxygenase component large chain - Pseudo
C:Species: Pseudomonas putida
C>Date: 15-Feb-1991 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C/Accession: A36516
R:Zylstra, G.J.; Gibson, D.T.
J. Biol. Chem. 264, 14940-14946, 1989
A:Title: Toluene degradation by Pseudomonas putida F1. Nucleotide sequence of the tod
A:Reference number: A36516; MUID:89359301; PMID:2670929
A:Accession: A36516
A:Residues: 1-450 <ZYLU>
A:Molecule type: DNA
A:Genetics:
A:Cross-References: GB:J04996; NID:g151600; PIDN:AAA26005.1; PID:g151601
A:Experimental source: strain F1
C:Genetics:
A:Gene: todC1
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [
C:Keywords: 2Fe-2S; aromatic hydrocarbon catabolism; metalloprotein; oxidoreductase;
F:86-134/Domain: Rieske [2Fe-2S] homology <RSK>
F:96,98,116,119/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 26.9%; Score 648.5; DB 1; Length 450;
Best Local Similarity 34.4%; Pred. No. 1,2e-43;
Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;

OY 20 IHGDELFQHELTIFARNMLFLTHDSLIPAPGDVYAKMGIDDEVYISRONDSIRAFIN 79
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 35 IYDDEDLQLELEVERFARSWLLGHETHIRKPGDYFTTYMKEDPVVYVVRQDAIAVFLN 94
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 80 VCHRRKTLVSEAGNAKGFVCSYHGWFSGNSELQSVPEKDLGESLNKKCLGLKE-- 137
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 95 QCHRRGRICRSADGAKNAKATTCSHGMAVDTAGNLINVPYAESFA-----CLNKKMS 148
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 138 --VARVESFHGFIYGCDFQDAPRLMDYLGAAYLLEPMF-KHSGLDELVGPFGKVVYKAN 194
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 149 PLKARVETVYKGLIFANNDENALDIDTYLGEAKFYNDHMLDRTEAGTEVYPIQKMWIPCN 208
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 195 WKPAPEFVODAHVGVGT-HASSLSRSG--ESIFSSLAGNALPREGAGLQMTSKYSGMG 251
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 209 WKFAEDQCSMDYHAGTTAHLGIIAGLPEDLELADLA--P-----KRGQYRAS 257
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 252 VLMDSYGVSHADLVPELMAF--GGAKOERLNKEIGDVR--ARIYRSHLCTVFPNNKSL 307
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 264 GFYVGDPNMLMIALINGPRVTSYTWGEPASAKAERLGSVERGSKLWVEM--TYFPCPSFL 321
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 308 TCSGVFVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVQRTTGPAGFESDNDNME 367
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 322 PGINTVTRWHRGRNEVEVMAFTYVDADAPDDIKEEPRROTLRFFSAGVFEEDDENNV 381
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
OY 368 TASQNGKKYQSDLSNLGFGEDVYGDVYGVGKSAIGETSYGFRATQAHVSSS 427
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
DB 382 EIQHILRCHAKSRPFNAEMSGQTVNDPDPYGRISNNVYSEEAARGLYAHMLRMTSP 441
| : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
```







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:46:47 ; Search time 29.1667 Seconds  
(without alignments)  
2051.298 Million cell updates/sec

Title: US-09-843-250-34  
2409  
Sequence: 1 MNYNKILVSESGLSQKHLI.....AEEFHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2409	100.0	449	21	AAAB12580
2	2404	99.8	449	21	AAAB12585
3	2404	99.8	449	21	AAAB12566
4	2404	99.8	449	21	AAAB12567
5	2404	99.8	449	21	AAAB12579
6	2403	99.8	449	21	AAAB12581
7	2403	99.8	449	21	AAAB12582
8	2402	99.7	449	21	AAAB12577
9	2402	99.7	449	21	AAAB12578
10	2402	99.7	449	21	AAAB12583

11	2399	99.6	449	21	AAAB12568
12	2397	99.5	448	21	AAAB12584
13	2394	99.4	449	21	AAAB12569
14	2346	97.4	449	21	AAAB12570
15	2318	96.2	449	21	AAAB12571
16	2292	95.1	449	21	AAAB12572
17	2219	92.1	449	21	AAAB12573
18	2189	90.9	447	21	AAAB12574
19	2052	85.2	447	21	AAAB12575
20	1968	81.7	451	21	AAAB12576
21	1093	45.4	452	21	AAAB29474
22	735	30.5	453	22	AAU34660
23	660.5	27.4	443	23	AAO17350
24	659.5	27.4	459	16	AAAB6729
25	621.5	25.8	459	16	AAAB6729
26	616	25.6	458	21	AAAB1990
27	613	25.4	458	21	AAAB1989
28	552.5	22.9	427	14	AAAB2085
29	375.5	15.6	385	22	AAU36099
30	375	15.6	490	22	AAAG2386
31	356	14.8	497	22	AAAB9227
32	336.5	14.0	424	22	AAU33601
33	239	9.9	435	22	AAAG2634
34	235	9.8	433	22	AAAG2635
35	235	9.8	433	22	AAAG2635
36	233	9.7	438	23	ABBO6798
37	224	9.3	446	19	AAAB69563
38	223	9.3	439	19	AAAB69562
39	196	8.1	99	23	ABP35375
40	186	7.7	426	23	ABG3163
41	182	7.6	1193	22	ABG18089
42	181	7.5	35	19	AAAB79018
43	173	7.2	35	19	AAAB79019
44	164	6.8	35	19	AAAB79020
45	126	5.2	405	19	AAAB80332

## ALIGNMENTS

RESULT 1	
AAAB12580	
AAAB12580 standard; Protein; 449 AA.	
AC	AAAB12580;
DT	09-NOV-2000 (first entry)
XX	
XX	Naphthalene dioxygenase mutant F352T protein sequence SEQ ID NO:34.
DE	
XX	
XX	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase: NDO;
KW	Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
OS	Pseudomonas sp.
OS	Synthetic.
PN	WO200037480-A1.
XX	
PD	29-JUN-2000.
XX	
PF	26-OCT-1999; 99WO-US25079.
XX	
PR	26-OCT-1999; 98US-0105575.
PA	(IOWA) UNIV IOWA RES FOUND.
XX	
PI	Parales R, Gibson D, Resnick S, Lee K;
XX	WPI; 2000-452174/39.
DR	N-PSDB; AAA65354.
XX	
PT	Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation

Claim 9, Page 129-131; 151pp; English.

PS The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 100.0%; Score 2409; DB 21; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 4,5e-225;  
 Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNVNNKILVSESGLSQKHLHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDVYATKMG 60  
 DB 1 MNVNNKILVSESGLSQKHLHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDVYATKMG 60  
 OY 61 IDEVIYSRQNDGSIKRAFLNVCRRHGRKTLVSVEGNMKGFCVSYHGWGFCSGNGLQSVPE 120  
 DB 61 IDEVIYSRQNDGSIKRAFLNVCRRHGRKTLVSVEGNMKGFCVSYHGWGFCSGNGLQSVPE 120  
 OY 121 KDLGSESLNKKCIGLKEVAVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180  
 DB 121 KDLGSESLNKKCIGLKEVAVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180  
 OY 181 ELVPGPKVYIKANMKAPAEENYGDVYHGWTHASSLRSEGSIFSSLAGNAALPPEGAGL 240  
 DB 181 ELVPGPKVYIKANMKAPAEENYGDVYHGWTHASSLRSEGSIFSSLAGNAALPPEGAGL 240  
 OY 241 QMTSKVSGMGLVMDYSGVHSADVLPELMAFGAKOERLNKEIGDVRATYRSHLCTV 300  
 DB 241 QMTSKVSGMGLVMDYSGVHSADVLPELMAFGAKOERLNKEIGDVRATYRSHLCTV 300  
 OY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIVEKMDPEDLKRLADSVORTGPAGFWES 360  
 DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIVEKMDPEDLKRLADSVORTGPAGFWES 360  
 OY 361 DDNDNMETYSQNGKKTQSRDLSLNGFGEEDYGAAYVYGVGKSAIGETSYRGFTFRAY 420  
 DB 361 DDNDNMETYSQNGKKTQSRDLSLNGFGEEDYGAAYVYGVGKSAIGETSYRGFTFRAY 420  
 OY 421 QAHVSSNNAAEFHASTWHTTELTKTTDR 449  
 DB 421 QAHVSSNNAAEFHASTWHTTELTKTTDR 449

RESULT 2  
 AAB12565  
 ID AAB12565 standard; Protein; 449 AA.  
 XX  
 AC AAB12565;

XX 09-NOV-2000 (first entry)

DE Naphthalene dioxygenase mutant F352V protein sequence SEQ ID NO:2.

KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS Pseudomonas sp.  
 OS Synthetic.

PN MO200037480-A1.

PD 29-JUN-2000.

PE 26-OCT-1999; 99MO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB: AAA65339.

PT Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,

PT pharmaceutical or rubber industry and for carrying out bioremediation

PS Claim 9, Page 58-60; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

SO Sequence 449 AA:

Query Match 99.8%; Score 2404; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 1.4e-224;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MNVNNKILVSESGLSQKHLHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDVYATKMG 60  
 DB 1 MNVNNKILVSESGLSQKHLHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDVYATKMG 60  
 OY 61 IDEVIYSRQNDGSIKRAFLNVCRRHGRKTLVSVEGNMKGFCVSYHGWGFCSGNGLQSVPE 120  
 DB 61 IDEVIYSRQNDGSIKRAFLNVCRRHGRKTLVSVEGNMKGFCVSYHGWGFCSGNGLQSVPE 120  
 OY 121 KDLGSESLNKKCIGLKEVAVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180  
 DB 121 KDLGSESLNKKCIGLKEVAVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180



XX 26-OCT-1998; 98US-0105575.  
PR  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
XX  
XX Parales R, Gibson D, Resnick S, Lee K;  
XX WPI: 2000-452174/39.  
DR N-PSDB: AAA65341.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
PS  
PS Claim 13; Page 100-102; 151pp: English.  
XX  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC benzothiofene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SQ  
Query Match 99.88; Score 2404; DB 21: Length 449;  
Best Local Similarity 99.88; Pred. No. 1,4e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNTNKKLIVSSSGLSOKRLIHGDELFQHELTTFARWMLFTHDSLIPACDDVYATMG 60  
DB 1 MNTNKKLIVSSSGLSOKRLIHGDELFQHELTTFARWMLFTHDSLIPACDDVYATMG 60  
QY 61 IDEVIVSRONGDSIRAFILNVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSNGELQSVPE 120  
DB 61 IDEVIVSRONGDSIRAFILNVCRRHGKTLVSVEAGNAKGFVCSYHGWGFGSNGELQSVPE 120  
QY 121 KDLYGESLNNKCLCLKEVAVRVSFHGFIYGCFFDOEAPPLMDYLGDAAVLEPMERHSGGL 180  
DB 121 KDLYGESLNNKCLCLKEVAVRVSFHGFIYGCFFDOEAPPLMDYLGDAAVLEPMERHSGGL 180  
QY 181 ELVGPPKVVYIKAMWKAPEAFVGDVAVHGMTHASSLSGSGSIFSSLAGNALPREAGGL 240  
DB 181 ELVGPPKVVYIKAMWKAPEAFVGDVAVHGMTHASSLSGSGSIFSSLAGNALPREAGGL 240  
QY 241 QMTSKYSYSGMGLVMDGYSGVADLPDELMAFGAKOERLKKETIGDVARARYSHLNTCV 300  
DB 241 QMTSKYSYSGMGLVMDGYSGVADLPDELMAFGAKOERLKKETIGDVARARYSHLNTCV 300  
QY 301 FPNNSMLTSGGVFVWMPIDANTEVWTYAIVEKMPEDLKRRLADSVQRTTGPAGFWS 360  
DB 301 FPNNSMLTSGGVFVWMPIDANTEVWTYAIVEKMPEDLKRRLADSVQRTTGPAGFWS 360  
QY 361 DDNDNMETASONGKRYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASONGKRYOSRSDLLSNLGFGEVDYGDVAYPGVYKSAIGETSYRGFYRAY 420

QY 421 QAHVSSNMWAEFEHASSTWHTELTKTTDR 449  
DB 421 QAHVSSNMWAEFEHASSTWHTELTKTTDR 449  
RESULT 5  
AAB12579  
ID AAB12579 standard; Protein: 449 AA.  
XX  
XX AAB12579;  
XX  
XX 09-NOV-2000 (first entry)  
XX  
XX Naphthalene dioxygenase mutant F352A protein sequence SEQ ID NO:33.  
DE  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
RW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
XX Pseudomonas sp.  
OS Synthetic.  
XX  
XX WO200037480-A1.  
XX  
XX 29-JUN-2000.  
XX  
XX 26-OCT-1999; 99WO-US25079.  
XX  
XX 26-OCT-1998; 98US-0105575.  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
XX  
XX Parales R, Gibson D, Resnick S, Lee K;  
XX WPI: 2000-452174/39.  
DR N-PSDB: AAA65353.  
XX  
XX Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
PT  
XX  
XX Claim 9; Page 128-129; 151pp: English.  
CC  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC benzothiofene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX Sequence 449 AA;  
SQ  
Query Match 99.88; Score 2404; DB 21: Length 449;  
Best Local Similarity 99.88; Pred. No. 1,4e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MMYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60
D 1 MMYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60
QY 61 IDEVIYSRONDGSIRAFVLCRRHRTLVSEVGNKAGFYCSYHMGFGSNGELQSVPE 120
D 61 IDEVIYSRONDGSIRAFVLCRRHRTLVSEVGNKAGFYCSYHMGFGSNGELQSVPE 120
QY 121 KDLYGSELNKKCGLEKEVARVESFHGFIYGCPOEARPLMDYLGDAWYLEPFRKHSGL 180
D 121 KDLYGSELNKKCGLEKEVARVESFHGFIYGCPOEARPLMDYLGDAWYLEPFRKHSGL 180
QY 181 ELVPGPKVYIKANMKAPAEVFGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPEGAGL 240
D 181 ELVPGPKVYIKANMKAPAEVFGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPEGAGL 240
QY 241 QMTSKYSGMGVLMDCYSGVHSADLYPELMAFGAGAKOERLNKEIGVRRARIYRSHLNCYV 300
D 241 QMTSKYSGMGVLMDCYSGVHSADLYPELMAFGAGAKOERLNKEIGVRRARIYRSHLNCYV 300
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVQRTGPAGFWES 360
D 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVQRTGPAGFWES 360
QY 361 DDNDNMETASONGKKYQSRDSDLNMLGFGEDYVDYVPGVYKSAIGETSYRGFYRAY 420
D 361 DDNDNMETASONGKKYQSRDSDLNMLGFGEDYVDYVPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNNAAEFEEHASTWHTTELKTTDR 449
D 421 QAHVSSNNAAEFEEHASTWHTTELKTTDR 449

RESULT 6
AAB12581
ID AAB12581 standard; Protein: 449 AA.
XX
AC AAB12581;
XX
DT 09-NOV-2000 (first entry)
XX
DE Naphthalene dioxygenase mutant F352L protein sequence SEQ ID NO:35.
XX
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX
OS Pseudomonas sp.
XX
OS Synthetic.
XX
PN MO200037480-A1.
XX
PD 29-JUN-2000.
XX
PE 26-OCT-1999; 99WO-US25079.
XX
PR 26-OCT-1998; 98US-0105575.
XX
PA (IOWA ) UNIV IOWA RES FOUND.
XX
PI Parales R, Gibson D, Resnick S, Lee K;
XX
DR MPI: 2000-452174/39.
XX
DR N-PSDB: AAA65355.
XX
PT Novel naphthalene dioxygenase mutant having a specific amino acid
substitution for preparing chiral diols for use in the polymer, resin,
pharmaceutical or rubber industry and for carrying out bioremediation
XX
PS Claim 9, Page 131-133; 151pp; English.
XX
```

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphanthrene a  
CC corresponding dihydro dihydro compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.8%; Score 2403; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred No. 1.7e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MMYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60
D 1 MMYNNKILVSESGLSQKHLIHGDEELFQHELKTIFARNMLFLTHDSLIPAPGDYVAKMG 60
QY 61 IDEVIYSRONDGSIRAFVLCRRHRTLVSEVGNKAGFYCSYHMGFGSNGELQSVPE 120
D 61 IDEVIYSRONDGSIRAFVLCRRHRTLVSEVGNKAGFYCSYHMGFGSNGELQSVPE 120
QY 121 KDLYGSELNKKCGLEKEVARVESFHGFIYGCPOEARPLMDYLGDAWYLEPFRKHSGL 180
D 121 KDLYGSELNKKCGLEKEVARVESFHGFIYGCPOEARPLMDYLGDAWYLEPFRKHSGL 180
QY 181 ELVPGPKVYIKANMKAPAEVFGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPEGAGL 240
D 181 ELVPGPKVYIKANMKAPAEVFGDAYHVGWTHASSLSRSGESIFSSLAGNAALPPEGAGL 240
QY 241 QMTSKYSGMGVLMDCYSGVHSADLYPELMAFGAGAKOERLNKEIGVRRARIYRSHLNCYV 300
D 241 QMTSKYSGMGVLMDCYSGVHSADLYPELMAFGAGAKOERLNKEIGVRRARIYRSHLNCYV 300
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVQRTGPAGFWES 360
D 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVQRTGPAGFWES 360
QY 361 DDNDNMETASONGKKYQSRDSDLNMLGFGEDYVDYVPGVYKSAIGETSYRGFYRAY 420
D 361 DDNDNMETASONGKKYQSRDSDLNMLGFGEDYVDYVPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNNAAEFEEHASTWHTTELKTTDR 449
D 421 QAHVSSNNAAEFEEHASTWHTTELKTTDR 449
```

```
RESULT 7
AAB12582
ID AAB12582 standard; Protein: 449 AA.
XX
AC AAB12582;
XX
DT 09-NOV-2000 (first entry)
XX
DE Naphthalene dioxygenase mutant F352I protein sequence SEQ ID NO:36.
XX
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
XX
```

inducible multi-component enzyme; alpha subunit; mutant; chiral diol; polymer; resin; pharmaceutical; rubber industry; bioremediation.

*Pseudomonas* sp. Synthetic.

WO200037480-A1.

29-JUN-2000.

26-OCT-1999; 99WO-US25079.

26-OCT-1998; 98US-0105575.

(IOWA ) UNIV IOWA RES FOUND.

Parales R, Gibson D, Resnick S, Lee K;

WPI: 2000-452174/39.

N-PSDB; AAA65356.

Novel naphthalene dioxygenase mutant having a specific amino acid substitution for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry and for carrying out bioremediation

Claim 9: Page 133-134; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphenanthrene or 3,4-dihydroxy-3,4-dihydronaphenanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzocyclophene, 9,10-dihydroanthracene, or 9,10-dihydronaphenanthrene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

	Query Match	99.8%	Score 2403	DB 21	Length 449
	Best Local Similarity	99.8%	Pred. No. 1.7e-224		
	Matches 448	Conservative	0	Mismatches 1	Indels 0
QY	1	MNYNKKILVSESGLSQKHLIHGDEDELFQHELTIFARNMLFLTHDSLIPAPGDIYATKMG	60		
Db	1	MNYNKKILVSESGLSQKHLIHGDEDELFQHELTIFARNMLFLTHDSLIPAPGDIYATKMG	60		
QY	61	IDEVYSQNODSITAFILNVCRRKGTILVSYERGNKKGVCYSYHGMGFSNGELQSVPE	120		
Db	61	IDEVYSQNODSITAFILNVCRRKGTILVSYERGNKKGVCYSYHGMGFSNGELQSVPE	120		
QY	121	KDLGESLKKKCLGLKEVARVSEFHGFIYGCPEQDEAPPLMDYLDGDAWYLEPFFKHSGL	180		
Db	121	KDLGESLKKKCLGLKEVARVSEFHGFIYGCPEQDEAPPLMDYLDGDAWYLEPFFKHSGL	180		
QY	181	ELVGPGRKVIYIANKKAPAEINLVGDIYHVGMTHASSLRSGESTFSSLAGNALPPEGAGL	240		
Db	181	ELVGPGRKVIYIANKKAPAEINLVGDIYHVGMTHASSLRSGESTFSSLAGNALPPEGAGL	240		

QY	241	QMTSKYSGMGVLMDGYSGVSHADLYPELMAFGAKQERLNKEIGDVARARYSHLNCY	300
Db	241	QMTSKYSGMGVLMDGYSGVSHADLYPELMAFGAKQERLNKEIGDVARARYSHLNCY	300
QY	301	FPNNSMLTCSGCFKFWKWNIDANTEVWYIAYVEKDMPEDLKRLADSVORTTGAGWES	360
Db	301	FPNNSMLTCSGCFKFWKWNIDANTEVWYIAYVEKDMPEDLKRLADSVORTTGAGWES	360
QY	361	DDNDMETSASONGKKKYOSRSDLLSNLGFGEADVYGDAYPGVWKSAGETSYRGFYRAY	420
Db	361	DDNDMETSASONGKKKYOSRSDLLSNLGFGEADVYGDAYPGVWKSAGETSYRGFYRAY	420
QY	421	QAHSSSWMAEPEHAASSTWHELTKTDDR	449
Db	421	QAHSSSWMAEPEHAASSTWHELTKTDDR	449

RESULT 8  
AAB12577  
ID AAB12577 standard; Protein; 449 AA.  
XX  
AC AAB12577:  
XX  
DT 09-NOV-2000 (first entry)  
DE Naphthalene dioxygenase alpha subunit protein sequence SEQ ID NO:26.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
RW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
PN MO200037480-A1.  
PD 29-JUN-2000.  
PF 26-OCT-1999; 99WO-US25079.  
PR 26-OCT-1998; 98US-0105575.  
PX (IOWA ) UNIV IOWA RES FOUNDED.  
PY Parales R, Glibson D, Resnick S, Lee K;  
PI WPI: 2000-452174/39.  
DR N-PSDB; AAA65351.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT -  
PS Disclosure; Page 118-120; 151pp; English.



CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC naphthalene dioxygenase (NDO) from the Pseudomonas sp. strain NCIB  
CC 9816-4, which is used in the exemplification of the present invention.  
XX

Sequence 449 AA:

Query Match 99.7%: Score 2402; DB 21; Length 449;  
Best Local Similarity 99.8%: Pred. No. 2.2e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYATKMG 60  
DB 1 MNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYATKMG 60  
QY 61 IDEVIYSRNDGSIKRAFLVNCRRHCKTLVSVKGNKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIYSRNDGSIKRAFLVNCRRHCKTLVSVKGNKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSLNKKCLGKEVARVESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180  
DB 121 KDLYGSLNKKCLGKEVARVESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180  
QY 181 ELVGPGRKVIYKANKKAPAEVFGDAYHWGTHASSLRSGESIFSSLAGNAALPPGAGL 240  
DB 181 ELVGPGRKVIYKANKKAPAEVFGDAYHWGTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVSHADLVEPELMARFGAGKOEERLKEIGVRRIRYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMDSYGVSHADLVEPELMARFGAGKOEERLKEIGVRRIRYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVORTGPAFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVORTGPAFWES 360  
QY 361 DDNDNMETASQNGKKYQSRDSDLSNMGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSRDSDLSNMGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNNAEFPHASSTWTELTCTTDR 449  
DB 421 QAHVSSNNAEFPHASSTWTELTCTTDR 449

RESULT 9

AAB12578 standard; Protein: 449 AA.

XX AAB12578;

DT 09-NOV-2000 (first entry)

DE Naphthalene dioxygenase mutant F352G protein sequence SEQ ID NO:32.

XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.

XX Pseudomonas sp.

OS Synthetic.

XX WO200037480-A1.

PD 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUNDED.

PI Parales R, Gibson D, Resnick S, Lee K;

XX

DR WPI: 2000-452174/39.

DR N-PSDB: AAA65352.

PT Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,

PT pharmaceutical or rubber industry and for carrying out bioremediation

PS Claim 9; Page 126-127; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidize an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzofuran,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzophenone, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

Sequence 449 AA:

Query Match 99.7%: Score 2402; DB 21; Length 449;  
Best Local Similarity 99.8%: Pred. No. 2.2e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYATKMG 60  
DB 1 MNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYATKMG 60  
QY 61 IDEVIYSRNDGSIKRAFLVNCRRHCKTLVSVKGNKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIYSRNDGSIKRAFLVNCRRHCKTLVSVKGNKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSLNKKCLGKEVARVESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180  
DB 121 KDLYGSLNKKCLGKEVARVESFHGFIYGCPEQAPPLMDYLGDAAWYLEPFRKHSGL 180  
QY 181 ELVGPGRKVIYKANKKAPAEVFGDAYHWGTHASSLRSGESIFSSLAGNAALPPGAGL 240  
DB 181 ELVGPGRKVIYKANKKAPAEVFGDAYHWGTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVSHADLVEPELMARFGAGKOEERLKEIGVRRIRYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMDSYGVSHADLVEPELMARFGAGKOEERLKEIGVRRIRYRSHLNCYV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVORTGPAFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVMTAIVEKMDPEDLKRLADSVORTGPAFWES 360  
QY 361 DDNDNMETASQNGKKYQSRDSDLSNMGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSRDSDLSNMGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNNAEFPHASSTWTELTCTTDR 449  
DB 421 QAHVSSNNAEFPHASSTWTELTCTTDR 449

RESULT 10

AA012583  
ID AAB12583 standard; Protein: 449 AA.  
XX  
AC AAB12583;  
XX  
DT 09-NOV-2000 (first entry)  
DE Naphthalene dioxygenase mutant F352W protein sequence SEQ ID NO:58.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO:  
inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
XX Synthetic.  
XX  
PN MO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
XX  
DR N-PSDB; AAA65376.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
substitution for preparing chiral diols for use in the polymer, resin,  
pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Example 7: Page 142-144; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
related complex (I) comprising several polypeptides which contain an  
alpha subunit that contains substituted amino acids at specific  
positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
exemplification of the present invention.  
XX  
XX  
SQ Sequence 449 AA:  
Query Match 99.7%; Score 2402; DB 21; Length 449;  
Best local Similarity 99.8%; Pred. No. 2.2e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNYNNKLVSESGISQKHLHGDDELFOHELTIFARNWMLFTHDSLTPABGDVYTAKG 60  
DB 1 MNNYNNKLVSESGISQKHLHGDDELFOHELTIFARNWMLFTHDSLTPABGDVYTAKG 60  
QY 61 IDEYIVSRQNDGSTRALFLNLCRHNGKTLVSEAGNAKGFVCSYIGWGGSGNGELQSVPE 120  
DB 61 IDEYIVSRQNDGSTRALFLNLCRHNGKTLVSEAGNAKGFVCSYIGWGGSGNGELQSVPE 120

DB 61 IDEYIVSRQNDGSTRALFLNLCRHNGKTLVSEAGNAKGFVCSYIGWGGSGNGELQSVPE 120  
QY 121 KDLYGESLNKKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMPKHSGL 180  
DB 121 KDLYGESLNKKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMPKHSGL 180  
QY 181 ELVGPBGKVIKAMKKADENFVDAYHVGWTHASSLSRSGESTFSSLAGNAALPPEGAGL 240  
DB 181 ELVGPBGKVIKAMKKADENFVDAYHVGWTHASSLSRSGESTFSSLAGNAALPPEGAGL 240  
QY 241 QMTSKYSGSGMGLMDGYSGVHSAADLVPMLAFGAKOERLKEIGDVARATYSHLNCY 300  
DB 241 QMTSKYSGSGMGLMDGYSGVHSAADLVPMLAFGAKOERLKEIGDVARATYSHLNCY 300  
QY 301 FPNNSMLTFCGCVFKWPNIDANTTEVMTYALVEKDMPEDLKRLADSVQRTGPAGFWES 360  
DB 301 FPNNSMLTFCGCVFKWPNIDANTTEVMTYALVEKDMPEDLKRLADSVQRTGPAGFWES 360  
QY 361 DDNDMMETASONGKKYQSRDLSNLGFGEDVYGDAAVYPGVWKSALGETSYRGFYRAY 420  
DB 361 DDNDMMETASONGKKYQSRDLSNLGFGEDVYGDAAVYPGVWKSALGETSYRGFYRAY 420  
QY 421 QAHVSSNMWAEFEHASSFWHTELTKTDR 449  
DB 421 QAHVSSNMWAEFEHASSFWHTELTKTDR 449  
RESULT 11  
AAB12568  
ID AAB12568 standard; Protein: 449 AA.  
XX  
AC AAB12568;  
XX  
DT 09-NOV-2000 (first entry)  
DE NDO related complex alpha subunit protein sequence SEQ ID NO:16.  
XX  
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO:  
inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
XX Synthetic.  
XX  
PN MO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
DR WPI: 2000-452174/39.  
XX  
DR N-PSDB; AAA65342.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
substitution for preparing chiral diols for use in the polymer, resin,  
pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Claim 13: Page 102-103; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
related complex (I) comprising several polypeptides which contain an  
alpha subunit that contains substituted amino acids at specific  
positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which

```
CC involves contacting them with naphthalene, biphenyl, phenanthrene,
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
CC involves contacting them with phenanthrene. The polypeptides and the
CC host cells are also used in bioremediation in which they oxidise an
CC aromatic compound such as indene, 1,2-dihydronaphthalene,
CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
CC dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a
CC corresponding dihydro dihydroxy compound. The polypeptides and the host
CC cells are useful for preparing chiral diols for use in the polymer,
CC resin, pharmaceutical or rubber industry. The present sequence represents
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
CC exemplification of the present invention.
XX
XX Sequence 449 AA:
SO
Query Match 99.6%; Score 2399; DB 21; Length 449;
Best Local Similarity 99.6%; Pred. No. 4.2e-224;
Matches 447; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MNNNNKILVSESGLSOKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYAKMG 60
DB 1 MNNNNKILVSESGLSOKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYAKMG 60
QY 61 IDEIVSRQNDGSIIRAFNLVNCRRHGRKTLVSEAGNAKGFCVSYHGFGSNGELQSVPE 120
DB 61 IDEIVSRQNDGSIIRAFNLVNCRRHGRKTLVSEAGNAKGFCVSYHGFGSNGELQSVPE 120
QY 121 KDLYGSLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180
DB 121 KDLYGSLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180
QY 181 ELVGPGRKVIYKANKKAPENFYGDAVHWGTHASSLRSGESFSSLAGNAALPPGAGL 240
DB 181 ELVGPGRKVIYKANKKAPENFYGDAVHWGTHASSLRSGESFSSLAGNAALPPGAGL 240
QY 241 QMTSKYSGMGVLMDSYSGVHSADLVPBELMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300
DB 241 QMTSKYSGMGVLMDSYSGVHSADLVPBELMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300
QY 301 FPNNSMLTCSGCVKWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTGPAFWES 360
DB 301 FPNNSMLTCSGCVKWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTGPAFWES 360
QY 361 DDNDNETASONGKTYOSRSDLSNLGFGEDYVGDVAVPGVYKSAIGTSRGRYRAY 420
DB 361 DDNDNETASONGKTYOSRSDLSNLGFGEDYVGDVAVPGVYKSAIGTSRGRYRAY 420
QY 421 QAHVSSNMAEFHASTWHTLTKTTDR 449
DB 421 QAHVSSNMAEFHASTWHTLTKTTDR 449
RESULT 12
AAB12584
ID AAB12584 standard; Protein; 448 AA.
XX
XX AAB12584;
XX
XX 09-NOV-2000 (first entry)
XX
XX Naphthalene dioxygenase mutant F352Y protein sequence SEQ ID NO:59.
XX
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
XX inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
XX polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX
XX Pseudomonas sp.
XX Synthetic.
XX
XX WO200037480-A1.
PN
```

```
XX
XX 29-JUN-2000.
PD
XX
XX 26-OCT-1999; 99MO-US25079.
PF
XX
XX 26-OCT-1998; 98US-0105575.
PR
XX
XX (IOWA ) UNIV IOWA RES FOUND.
PA
XX
XX Parales R, Gibson D, Resnick S, Lee K:
XX
XX WPI: 2000-452174/39.
DR
XX
XX N-PSDB: AAA65377.
PT
XX
XX Novel naphthalene dioxygenase mutant having a specific amino acid
PT substitution for preparing chiral diols for use in the polymer, resin,
PT pharmaceutical or rubber industry and for carrying out bioremediation
PT .
XX
XX Example 7; Page 144-145; 151pp; English.
XX
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO
XX related complex (1) comprising several polypeptides which contain an
XX alpha subunit that contains substituted amino acids at specific
XX positions. The polypeptides and host cells are useful for preparing
XX (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
XX dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-
XX dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
XX involves contacting them with naphthalene, biphenyl, phenanthrene,
XX indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
XX the host cells are also useful for preparing 1,2-dihydroxy-1,2-
XX dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
XX involves contacting them with phenanthrene. The polypeptides and the
XX host cells are also used in bioremediation in which they oxidise an
XX aromatic compound such as indene, 1,2-dihydronaphthalene,
XX benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,
XX acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
XX dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a
XX corresponding dihydro dihydroxy compound. The polypeptides and the host
XX cells are useful for preparing chiral diols for use in the polymer,
XX resin, pharmaceutical or rubber industry. The present sequence represents
XX a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
XX exemplification of the present invention.
XX
XX Sequence 448 AA:
SO
Query Match 99.5%; Score 2397; DB 21; Length 448;
Best Local Similarity 99.8%; Pred. No. 6.6e-224;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MNNNNKILVSESGLSOKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYAKMG 60
DB 1 MNNNNKILVSESGLSOKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAPGDIYAKMG 60
QY 61 IDEIVSRQNDGSIIRAFNLVNCRRHGRKTLVSEAGNAKGFCVSYHGFGSNGELQSVPE 120
DB 61 IDEIVSRQNDGSIIRAFNLVNCRRHGRKTLVSEAGNAKGFCVSYHGFGSNGELQSVPE 120
QY 121 KDLYGSLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180
DB 121 KDLYGSLNKKCLGKLEVARVESFHGFIYGCFOEAPPLMDYLGDAWYLEPMFKHSGGL 180
QY 181 ELVGPGRKVIYKANKKAPENFYGDAVHWGTHASSLRSGESFSSLAGNAALPPGAGL 240
DB 181 ELVGPGRKVIYKANKKAPENFYGDAVHWGTHASSLRSGESFSSLAGNAALPPGAGL 240
QY 241 QMTSKYSGMGVLMDSYSGVHSADLVPBELMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300
DB 241 QMTSKYSGMGVLMDSYSGVHSADLVPBELMAFGAKOERLNKEIGDVRARIYRSHLNCYV 300
QY 301 FPNNSMLTCSGCVKWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTGPAFWES 360
DB 301 FPNNSMLTCSGCVKWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTGPAFWES 360
```

0y	361	DNDNDNETSONKCKKQSRSDLLSNLGFEDDYGAIVYGVCKSAIGETSYGCFRAY	420
Db	361	DDNDNDNETASQNKCKKQSRSDLLSNLGFEDDYGAIVYGVKSAIGETSYGCFRAY	420
Qy	421	QAHVSSSNMAEFHASTWHTLTCTTD	448
Db	421	QAHVSSSNMAEFHASTWHTLTCTTD	448
RESULT 13			
AAB12569			
ID	AAB12569	standard; Protein: 449 AA.	
XX			
AC	AAB12569;		
XX			
DT	09-NOV-2000	(first entry)	
XX			
DE	NDO related complex alpha subunit protein sequence SEQ ID NO:17.		
XX			
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;		
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
XX			
OS	Pseudomonas sp.		
OS	Synthetic.		
XX			
PN	WO200037480-A1.		
XX			
PD	29-JUN-2000.		
XX			
PF	26-OCT-1999;	99WO-US25079.	
XX			
PR	26-OCT-1998;	98US-0105575.	
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
XX			
PI	Parales R, Gibson D, Resnick S, Lee K;		
XX			
DR	WPI: 2000-452174/39.		
XX			
PS	N-PSDB; AAA65343.		
XX			
XX	Claim 13; Page 104-105; 151pp; English.		
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO		
CC	related complex (I) comprising several polypeptides which contain an		
CC	alpha subunit that contains substituted amino acids at specific		
CC	positions. The polypeptides and host cells are useful for preparing		
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-		
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-		
CC	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which		
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,		
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and		
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-		
CC	dihydronaphthanene or 3,4-dihydroxy-3,4-dihydronaphthanene which		
CC	involves contacting them with phenanthrene. The polypeptides and the		
CC	host cells are also used in bioremediation in which they oxidise an		
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,		
CC	benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,		
CC	acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,		
CC	dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthanene a		
CC	corresponding dihydro dihydroxy compound. The polypeptides and the host		
CC	cells are useful for preparing chiral diols for use in the polymer,		
CC	resin, pharmaceutical or rubber industry. The present sequence represents		
CC	a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the		
CC	exemplification of the present invention.		
XX			
XX	Sequence 449 AA:		

Query Match	99.4%	Score 2394	DB 21	Length 449
Best Local Similarity	99.3%	Pred. No. 1.3e-223		
Matches 446	Conservative 1	Mismatches 2	Indels 0	Gaps 0
Oy	1	MNYNNKILVSESGLSQKHLIHDEDELPFOHELKTIPARNWMLFLTHDSLLTPAGDVTYAKMG	60	
Db	1	MNYNNKILVSESGLSQKHLIHDEDELPFOHELKTIPARNWMLFLTHDSLLTPAGDVTYAKMG	60	
Oy	61	IDEVIVSRQNDGSIIRAFILNVCRHKGTLVSVYAGNAKGFVCSYHGWFGSGNGELQSVPE	120	
Db	61	IDEVIVSRQNDGSIIRAFILNVCRHKGTLVSVYAGNAKGFVCSYHGWFGSGNGELQSVPE	120	
Oy	121	KDLVSESLNKKCTGLKEVARVESFHGFTYGGCFDQAPRLMDYLDAAAYLEPMKHHSGGL	180	
Db	121	KDLVSESLNKKCTGLKEVARVESFHGFTYGGCFDQAPRLMDYLDAAAYLEPMKHHSGGL	180	
Oy	181	ELVGPGRGVYIKANKAKAPAEFVGDYAVHVGWTHASSLSRSGESIFSSLAGNALPPEGAGL	240	
Db	181	ELVGPGRGVYIKANKAKAPAEFVGDYAVHVGWTHASSLSRSGESIFSSLAGNALPPEGAGL	240	
Oy	241	QMTSKYSGSGMGVLMGDYSGVSHADLVPELMAFGAKOERLNKEIGDVARARIYRSHLNCV	300	
Db	241	QMTSKYSGSGMGVLMGDYSGVSHADLVPELMAFGAKOERLNKEIGDVARARIYRSHLNCV	300	
Oy	301	FPNNMFLCSCGFVKWNPIDANTPEVWTVYATVEXKMPEDLKRRLADSVQRTTGPAGFMES	360	
Db	301	FPNNMFLCSCGFVKWNPIDANTPEVWTVYATVEXKMPEDLKRRLADSVQRTTGPAGFMES	360	
Oy	361	DDNDMMETASONGKKGKQSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFFRAY	420	
Db	361	DDNDMMETASONGKKGKQSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFFRAY	420	
Oy	421	QAHVSSSWMAEPFENASSTWHTELTKTDR	449	
Db	421	RAHVSSSWMAEPFENASSTWHTELTKTDR	449	







Matches	165:	Conservative	64:	Mismatches	174:	Indels	50:	Gaps
QY	20	INDELEFQHELKTIFARWMLFETHDSLIAPEDYATAKMGIDEIVSRONGSITAFIN	79					
Db	24	1YTDPPIDYLETERRIGRCWLFELAHSEIQPKPDEDFNTYMGEDAVVYVQKGSISAKF	83					
QY	80	VQRHREKTLVSVEAGNAKKGVCYHSGMGSGNSELOSPFENKLVGESINKKCLGKEVA	139					
Db	84	QCHRRMRVSYACGCTRAFTCTCYHGMSTGINELIDVLEPRATYQGLCKSHWGLNEVP	143					
QY	140	RYSEFNGFYGFCDQAPPLMDYLGDAWYLEPMF-KHSGLELYGPRGKVVIKANMKAP	198					
Db	144	CYESYGLGLFGMNDTSAPGLRDYLGDIAWYLDGMIDLRREGTEITVGQKQWVINCWKRP	203					
QY	199	AENFYGDAYHVGCTHASSL-----NSGESIFSSLGNAALP-----	234					
Db	204	ASQFASDOYHALFSSASAVQVLGAKDKGSDPKRLGD-----GQTRAPWEETAKDALQEG	256					
QY	235	--PEGAGLQMTSYSGSGMGYLMDYSGVSHADLVPELMAFGAKQGRLNKEIGDVARLY	292					
Db	257	QDGHSGEFFTEKPRDANVWY--DGAVSYYRETYAE-----AEQRLGEVYALRL	303					
QY	293	RSHLNCTVEPNNSMLTQSGVFKVMPIDANTTEWMTYAIVEKMPEDLKRRLADSYORTL	352					
Db	304	AGHNH--IEPTLSMLNGTATLRYWHRPQDQYVMAFCITDKAASDEVKAAPENSATRAP	361					
QY	353	GRAGFESDNDNMETASQNGKKYGRSDLSLNGFGEIDYGDVAYPGVYKSAIGETS	412					
Db	362	GRAGFEQDSEWCEIQKLKGHRARNRKLCEMLGIGOEKRRDDIGPIT-NYIPSETA	420					
QY	413	YRGFYRAYQAHVSSSNMAEFEEHASSYWHLETK	445					
Db	421	ARGMYQRMADLLSSSEWQEVLDKTAAYQDEYVK	453					
RESULT 2								
US-09-738-626-6140								
: Sequence 6140, Application US/09738626								
: Publication No. US20020197605A1								
: GENERAL INFORMATION:								
: APPLICANT: NAKAGAMA, SATOSHI								
: APPLICANT: MIZOGUCHI, HIROSHI								
: APPLICANT: ANDO, SEIKO								
: APPLICANT: HAYASHI, MIKIRO								
: APPLICANT: OCHIAI, KEIKO								
: APPLICANT: YOKOI, HARUHIKO								
: APPLICANT: YATEISHI, NAOKO								
: APPLICANT: SENOH, AKIHIRO								
: APPLICANT: IKEDA, MASATO								
: APPLICANT: OZAKI, AKIO								
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES								
: FILE REFERENCE: 249-125								
: CURRENT APPLICATION NUMBER: US/09/738, 626								
: CURRENT FILING DATE: 2000-12-18								
: PRIOR APPLICATION NUMBER: JP 99/377484								
: PRIOR FILING DATE: 1999-12-16								
: PRIOR APPLICATION NUMBER: JP 00/159162								
: PRIOR FILING DATE: 2000-04-07								
: PRIOR APPLICATION NUMBER: JP 00/280988								
: PRIOR FILING DATE: 2000-08-03								
: NUMBER OF SEQ ID NOS: 7059								
: SOFTWARE: PatentIn ver. 3.0								
: SEQ ID NO 61								

Query Match	15.7%	Score 377	DB 9	Length 490
Best Local Similarity	26.3%	Pred. No. 1.4e-25		
Matches 118	Conservative 66	Mismatches 215	Indels 50	Gaps 15

[illegible]

```

RESULT 3
US-09-815-242-11692
Sequence 11692, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Pokeryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11692
LENGTH: 385
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-815-242-11692

```

Query Match 15.68; Score 376.5; DB 10; Length 385;



[illegible]

```

Db      24 ELHRRELHEIFDSDWLYAAHLSLELRPEGDFITRDVGGNNLITQRRADGEPAVYLNACAH 83
Oy      85 GKTIVSVEAGNANGFVCSYSHWMGFGSNGELOSVPEKDLYGESLNKCC--LGLKEVAVY 14
Db      84 GAKCAERQGNRSQRCPTPHGWYDHSGLSLP-DRKAYHA--GGCHPELSLTQYKHA 14
Oy      142 ESFNGFIYFGDFDEAPPLNDYLDGAAMYLEPMFKHS-GGLELVGPCKVYIKANMKPAE 20
Db      141 -VYNNFLFIHYGARQASLETFYLGQADYIDLICDQSEAELEITPGGFENSIKANMKLAE 19
Oy      201 NFVQDAVHGVTAASSLRGESEIFSSLAGNAALPREGGLQMTKYSSGKGYLMDYSGV 26
Db      200 NGV-DAVHPEFAHKRYLEYLNTL-----GTDPESHRHGRG- EALGNHALI 24
Oy      261 HS-----ADLVEP-LMAFGAKOERLNKIEIGDVAR-IYRSHLNCVFPN--- 30
Db      245 ISGPPSTGRPLATWSPLEFPALKPSTIAAFAEFLVERFEQQAEDIAHTNKSLETFPLVY 30
Oy      304 NSMLTCSGVFVKVNPIDANTTEWYTAIYEKMPEDLKRRLADSVQRTLGPAGFWESDDN 36
Db      305 NDILGLN--IRSFPLAADSVSYTWVAGAFNDETRERAAINILNLFIPGGFGCPDDV 36
Oy      364 DNMETASQNGKTKYQSRDLSLNLGFGEDYGDVAVPGVWKSAGTSTYGRFYAV 42
Db      363 FILESCQ---RAYAH-----AALGYSDFSRG---MGPATRRHVHDEQNRGFWREM 406

RESULT 5
US-09-776-490-19
; Sequence 19, Application US/09776490
; Patent No. US20010012886A1
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
;              John, Gurmukh S.
;              Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
;                   CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: NO. US20010012886A1th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/776,490
; FILING DATE: 02-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/810,009
; FILING DATE: 04-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS: <unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;

```

SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-490-19

Query Match 7.7%; Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%; Pred. No. 3.2e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAKGFCVSYHGMGFGSNGE 113  
|||||  
Db 1 NVCRRHGRKTLVSEAGNAKGFCVSYHGMGFGSNGK 35

RESULT 6  
US-09-776-491-19  
Sequence 19, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-491-19

Query Match 7.7%; Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%; Pred. No. 3.2e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAKGFCVSYHGMGFGSNGE 113  
|||||  
Db 1 NVCRRHGRKTLVSEAGNAKGFCVSYHGMGFGSNGK 35

RESULT 7  
US-09-776-490-20  
Sequence 20, Application US/09776490  
Patent No. US20010012886A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-490-20

Query Match 7.2%; Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%; Pred. No. 3.1e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAKGFCVSYHGMGFGSNGE 113  
|||||  
Db 1 NVCRRHGRKTLVSEAGNAKGFCVSYHGMGFGSNGK 35

RESULT 8  
US-09-776-491-20  
Sequence 20, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-491-20

Query Match 7.2%; Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%; Pred. No. 3.1e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCRRHKGTLVSEAGNAGKFCVSGHGWGSGNGE 113  
DB 1 NVCRRHKGTLVSEAGNAGKFCVSGHGWGSGNGK 35

RESULT 9  
US-09-776-490-21  
Sequence 21, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-490-21

Query Match 7.0%; Score 168; DB 10; Length 35;  
Best Local Similarity 80.0%; Pred. No. 1e-08;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCRRHKGTLVSEAGNAGKFCVSGHGWGSGNGE 113  
DB 1 NVCRRHKGTLVSEAGNAGKFCVSGHGWGSGNGK 35

RESULT 10  
US-09-776-491-21  
Sequence 21, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-491-21

Query Match 7.0%; Score 168; DB 10; Length 35;  
Best Local Similarity 80.0%; Pred. No. 1e-08;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;



CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-491-12

Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0014;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAGKGCSTYHGMAFYDSAG 112  
DB 1 NVCRRHGRKTLVSEAGNAGKGCSTYHGMAFYDSAG 34

RESULT 14  
US-09-776-490-14  
Sequence 14, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-776-490-14

Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0018;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTLVSEAGNAGKGCSTYHGMAFYDSAG 112  
DB 1 NVCRRHGRKTLVSEAGNAGKGCSTYHGMAFYDSAG 34

RESULT 15  
US-09-776-491-14  
Sequence 14, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-776-491-14

Query Match 4.5%: Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0018;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHGRKTLVSVENAGNAKGFCVSYHGWGFGSNG 112  
| | | | | : | | | | | : | |  
Db 1 NCRHRGMRICRADGNAKSPCTCSYHGWAYDSAG 34

Search completed: January 27, 2003, 09:20:57  
Job time : 8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:56:07 ; Search time 10.1667 Seconds  
(without alignments)  
1299.432 Million cell updates/sec

Title: US-09-843-250-35  
Perfect score: 2408  
Sequence: 1 MNYNKILVSEGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	224	9.3	446	US-09-004-393B-4	Sequence 4, Appli
2	223	9.3	439	US-09-004-393B-2	Sequence 2, Appli
3	185	7.7	35	US-08-810-009-19	Sequence 19, Appli
4	174	7.2	35	US-08-810-009-20	Sequence 20, Appli
5	168	7.0	35	US-08-810-009-21	Sequence 21, Appli
6	113.5	4.7	379	US-09-028-934-36	Sequence 36, Appli
7	110	4.6	35	US-08-810-009-12	Sequence 12, Appli
8	109	4.5	35	US-08-810-009-14	Sequence 14, Appli
9	108	4.5	35	US-08-810-009-13	Sequence 13, Appli
10	108	4.5	35	US-08-810-009-15	Sequence 15, Appli
11	108	4.5	622	US-09-311-626B-4	Sequence 4, Appli
12	106	4.4	35	US-08-810-009-9	Sequence 9, Appli
13	104	4.3	35	US-08-810-009-18	Sequence 18, Appli
14	103.5	4.3	432	US-08-809-326A-16	Sequence 16, Appli
15	103.5	4.3	649	US-08-809-326A-15	Sequence 15, Appli
16	103	4.2	35	US-08-810-009-11	Sequence 11, Appli
17	102	4.2	35	US-08-810-009-8	Sequence 8, Appli
18	102	4.2	35	US-08-810-009-10	Sequence 10, Appli
19	101	4.2	17	US-08-810-009-44	Sequence 44, Appli
20	101	4.2	17	US-08-810-009-45	Sequence 45, Appli
21	100	4.1	35	US-08-810-009-17	Sequence 17, Appli
22	99.5	4.1	256	US-09-325-932A-57	Sequence 57, Appli
23	95	3.9	17	US-08-810-009-46	Sequence 46, Appli
24	94.5	3.9	350	US-08-828-922-1	Sequence 1, Appli
25	94	3.9	560	US-08-814-052-6	Sequence 6, Appli
26	94	3.9	560	US-08-812-829-6	Sequence 6, Appli
27	93.5	3.9	446	US-07-627-539G-7	Sequence 7, Appli

28	93.5	3.9	468	US-07-627-539G-2	Sequence 2, Appli
29	93	3.9	35	US-08-810-009-16	Sequence 16, Appli
30	93	3.9	421	US-09-239-303-2	Sequence 2, Appli
31	93	3.9	563	US-09-134-001C-4800	Sequence 4800, Ap
32	93	3.9	5588	US-09-036-981A-6	Sequence 6, Appli
33	93	3.9	5588	US-09-370-700-6	Sequence 6, Appli
34	92.5	3.8	774	US-09-276-400-7	Sequence 7, Appli
35	92.5	3.8	774	US-09-448-076-7	Sequence 7, Appli
36	92.5	3.8	774	US-09-702-572-7	Sequence 7, Appli
37	92	3.8	1087	US-08-570-311-8	Sequence 8, Appli
38	92	3.8	1087	US-08-353-485-8	Sequence 8, Appli
39	92	3.8	1358	US-08-570-311-27	Sequence 27, Appli
40	92	3.8	1693	US-08-478-507-7	Sequence 7, Appli
41	92	3.8	1693	US-09-128-275A-7	Sequence 7, Appli
42	92	3.8	1693	US-09-553-427-7	Sequence 7, Appli
43	91.5	3.8	525	US-08-348-891A-2	Sequence 2, Appli
44	91.5	3.8	525	US-08-905-817-2	Sequence 2, Appli
45	90.5	3.8	345	US-09-362-123A-2	Sequence 2, Appli

#### ALIGNMENTS

```
RESULT 1
US-09-004-393B-4
; Sequence 4, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
; APPLICANT: Rathinasabapathi, Balu
; APPLICANT: Burnett, Michael
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; TITLE OF INVENTION: Plants Transformed Therewith
; FILE REFERENCE: US-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Beta vulgaris
US-09-004-393B-4

Query Match          9.3%; Score 224; DB 4; Length 446;
Best Local Similarity 27.4%; Pred. No. 6.7e-13;
Matches 58; Conservative 37; Mismatches 97; Indels 20; Gaps 6;

QY 11 EESGLSQKHLIHDEDELFQHEKLTIFARNWLFTHDSLIPAPEDDYTAKKGIDEVIYSRON 70
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
DB 99 EALTPPSTWYEPAPFYSHELRIFFYKQWAGVSEQVEKNQYFGSLGNEYLVSRDG 158
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
QY 71 DGSIRAPLVNCHRRKGTIVSEAGNAKFCVSYHGFGFSGNELDSVPREKDYGSLSNK 130
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
DB 159 QEELIARFHVCHHRA-SILACSGKSKSCVPCYHGWVYGLDLSLAKA--SKATETQNDLP 215
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
QY 131 KCLGKEVARVESFGEFTYGCDFQDEAPRLMD---YLSDA-----AWYLEPMFKHSGGLE 181
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
DB 216 KELGIAPL-KVAEWGPFILISIDRSIDANADVGTEWIGSARDVKAHARDPNLKFTHRSE 274
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
QY 182 LVGPGRKVVYIKAKAPAEFYGDAYHYVQWTH 213
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |
DB 275 F-----PMECNMKVFCNDYLDSSYHVPYAH 299
   | | : : : : | | : : : : | | : : : : | | : : : : | | : : : : | |

RESULT 2
US-09-004-393B-2
; Sequence 2, Application US/09004393B
; Patent No. 6310271
; GENERAL INFORMATION:
; APPLICANT: Hanson D., Andrew
```

RESULT 5  
US-08-810-009-21



```
; Sequence 21, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-21

Query Match          7.0%; Score 168; DB 4; Length 35;
Best Local Similarity 80.0%; Pred. No. 2.8e-09;
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NYCRRHKGTLVSEAGNKGFCYSHGWFSGSNGE 113
Db 1 NYCRRHKGTLVDAEGNAGPCYGHGWSGNGK 35

RESULT 6
US-09-028-934-36
; Sequence 36, Application US/09028934
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Helinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrirolnitrin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 611760artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-028-934-36

Query Match          4.7%; Score 113.5; DB 3; Length 379;
Best Local Similarity 20.3%; Pred. No. 0.016;
Matches 95; Conservative 54; Mismatches 149; Indels 171; Gaps 26;
```

```
QY 20 IHGDEELFQ-----HEKTIIFARNW-LFTHDSLIAPDQDYTAKKGIDEVIYSRQDGS 73
Db 5 IHQEPERIQASGVNDLTQTOTAAWYVAMRSDALRGKPV--VAIKLFGQPLVAMRGGR 61

QY 74 IRAFLNVCRRHCKTLVSEAGNKGFCV-----CSYHGMGFGSNGELQSVPEFKDLYGS 127
Db 62 PYVMERYCSHLGASL-----AKGKYVEGCIOCPPHNMRIIDSTGCASHVP-----GHS 108

QY 128 LNKKCLGLEKVARVESF-----HGFIYGCEDQAP-----PLMDYLGDAWY 169
Db 109 -----TEVRLEPIPTARQSVYPVMERYGFVWVGTRAPLPLEMPAESESEH 160

QY 170 LEPMEKHSGLGLVCPGPGVYIKAMKAPAEENFGDAYIVGWTTHASSLSGE-SIF---S 225
Db 161 QSLRFAYETTTSVL-----RII-----ENFY-DAQHAAPVHQLPISAFELKLFDESS 206

QY 226 SLAGNAALPPE-----GAGLQW-TSKYSGMGVL-----WPGYSGVHSADLV 266
Db 207 PPPGOEALARDGAMFGAGIDFHYDRIFGFLGVISRTGLGSMBSMQLHFPGYRP----- 259

QY 267 PELMAFGAKOERLNEIDVBARIRYRSHLNTVFPNNSMLTCSGVFKWNPIDANTTEY 326
Db 260 -----GCIMTVSLD---GDVKYRL---LQCV-----TPDKRETYW 289

QY 327 WTYAIVKQMPEDLKRRLADSV---ORTLGPAFG---WESDNDNMETASONGKKYQSR 379
Db 290 HHLLAIKKG--DGVVSANFLIYGLOTWAAGAAGYDAIVNMSMKAD-----GGGAFSKY 340

QY 380 DSDLSNLGFGEDYGDAYPCVYVKSALGERSYRGFRAYQAHVSSSN 428
Db 341 DDLIL-----KYRAFRRWVKVALEN 362

RESULT 7
US-08-810-009-12
; Sequence 12, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
```

APPLICANT: Johal, Gurmukh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-12

Query Match 4.6%; Score 110; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0009;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NVCRHKGKTLVSVAGNAKGFVCSYHGWFSGSNG 112  
| | | | | : | | | | | : | |  
Db 1 NQCRHGRMCRADGNAKAPTCSYHGWAYDVTAG 34

RESULT 8  
US-08-810-009-14  
Sequence 14, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-14

Query Match 4.5%; Score 109; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0011;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHKGKTLVSVAGNAKGFVCSYHGWFSGSNG 112  
| | | | | : | | | | | : | |  
Db 1 NQCRHGRMCRADGNAKAPTCSYHGWAYDSAG 34

RESULT 9  
US-08-810-009-13  
Sequence 13, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-13

Query Match 4.5%; Score 108; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0014;

	Matches	18;	Conservative	4;	Mismatches	12;	Indels	0;	Gaps
QY	79	NVCRHGRKTLVSEAGNAKGFVCSYHGNGFSGS	112		:			:	:
Db	1	NQCRHGRMTRCRSDAGNAKAPFCSYHGNAIYDAG	34						

RESULT 10  
US-08-810-009-15  
; Sequence 15, Application US/08810009  
; Patent No. 6211437

Query Match	4.5%	Score 108	DB 4	Length 35
Best Local Similarity	50.0%	Pred. No. 0.0014		
Matches 17	Conservative 4	Mismatches 13	Indels 0	Gaps 0
QY	79	NVCRHRGKTIYVSEGNNAKGPVCSITHMGFGSNG	112	
		:           :		
Db	1	NOCRHRGRIRICPADGNNAKSPFCSTYHGVAAYDVG	34	

RESULT 11  
 US-09-311-626B-4  
 ; Sequence 4, Application US/09311626B  
 ; Patent No. 6399347  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jorgensen, Per Lina  
 ; APPLICANT: Schnorr, Kirk  
 ; APPLICANT: Andersen, Leene No. 6399347boe  
 ; APPLICANT: Schuelein, Martin  
 ; APPLICANT: Outtrup, Helle  
 ; TITLE OF INVENTION: No. 6399347el Rhamnogalacturonan Hydrolases  
 ; FILE REFERENCE: 5572.204-US  
 ; CURRENT APPLICATION NUMBER: US/09/311,626B

```

: CURRENT FILING DATE: 1999-05-13
: PRIOR APPLICATION NUMBER: 0608/98
: PRIOR FILING DATE: 1998-05-01
: PRIOR APPLICATION NUMBER: 60/084,358
: PRIOR FILING DATE: 1998-05-05
: NUMBER OF SEQ ID NOS: 61
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 622
: TYPE: PRN
: ORGANISM: Bacillus licheniformis
:
:
US-09-311-626B-4

```

```

12 RESULT 12
13 US-08-810-009-9
14 : Sequence 9, Application US/08810009
15 : Patent No. 6211437
16 : GENERAL INFORMATION:
17 : APPLICANT: Briggs, Steven P.
18 : APPLICANT: Johal, Gurmukh S.
19 : APPLICANT: Gray, John
20 : TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
21 : TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
22 : NUMBER OF SEQUENCES: 65
23 : CORRESPONDENCE ADDRESS:
24 : ADDRESSEE: BELL, SLEUTZER, PARK & GIBSON
25 : STREET: P.O. Drawer 34009
26 : CITY: Charlotte
27 : STATE: No. 6211437/cb Carolina
28 : COUNTRY: USA
29 : ZIP: 28234
30 : COMPUTER READABLE FORM:
31 : MEDIUM TYPE: Floppy disk
32 : COMPUTER: IBM PC compatible
33 : OPERATING SYSTEM: PC-DOS/MS-DOS
34 : SOFTWARE: PatentIn Release #1.0, Version #1.30
35 : CURRENT APPLICATION DATA:
36 : APPLICATION NUMBER: US/08/810,009
37 : FILING DATE: 04-MAR-1997
38 : CLASSIFICATION: 800

```

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-9

Query Match 4.48; Score 106; DB 4; Length 35;  
Best Local Similarity 51.48; Pred. No. 0.0022;  
Matches 18; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

OY 79 NVCRHGRKTVSVAGNAKGFVCSYHGMFGSGNGE 113  
DB 1 NSCRHGRICRSDAGNAKAPTCTSHGAYADIAGK 35

RESULT 13  
US-08-810-009-18  
Sequence 18, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-18

Query Match 4.38; Score 104; DB 4; Length 35;  
Best Local Similarity 52.96; Pred. No. 0.0033;  
Matches 18; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

OY 79 NVCRHGRKTVSVAGNAKGFVCSYHGMFGSGNG 112  
DB 1 NSCRHGRALLCPFSGNKGKHFVCRKHGMSYDSSG 34

RESULT 14  
US-08-809-326A-16  
Sequence 16, Application US/08809326A  
Patent No. 6165478  
GENERAL INFORMATION:  
APPLICANT: Izutsu, Hiroshi  
APPLICANT: Obara, Kazuhiko  
APPLICANT: Matsumoto, Akira  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES,  
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAs, TRANSFORMA  
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM  
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106009/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106010/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106011/95  
FILING DATE: 28-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Charles E.  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 7426-043-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:54:32 : Search time 25.8333 Seconds  
(without alignments)  
3581.232 Million cell updates/sec

Title: US-09-843-250-35  
Perfect score: 2408  
Sequence: 1 MNNYKILVSESGLSQKHLI.....AEFEHASTWHTELTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp protent:\*  
12: sp.virus:\*  
13: sp.yeast:\*  
14: sp.unclassified:\*  
15: sp.virus:\*  
16: sp.bacteriap:\*  
17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2332	96.8	449	2	Q938R9
2	2292	95.2	449	2	Q52142
3	2208	91.7	449	2	Q92173
4	2193	91.1	447	2	Q915D2
5	2189	90.9	447	2	Q52382
6	2072	86.0	447	2	Q8VUD4
7	2054	85.3	447	2	P95564
8	2025	84.1	447	2	Q8RT14
9	1967.5	81.7	451	2	Q45695
10	1490	61.9	277	2	Q9ETK2
11	1488	61.8	277	2	Q9F5S4
12	1486	61.7	277	2	Q9F5T3
13	1484	61.6	277	2	Q9F5S8
14	1483	61.6	277	2	Q9F5S9
15	1451	60.3	277	2	Q9F5S3
16	1405	58.3	277	2	Q9EUC7

17	1403	58.3	277	2	Q9F5T8	Q9F5T8 pseudomonas
18	1402	58.2	277	2	Q9F5T5	Q9F5T5 pseudomonas
19	1399	58.1	277	2	Q9ETV3	Q9ETV3 pseudomonas
20	1399	58.1	277	2	Q9F5S5	Q9F5S5 pseudomonas
21	1396	58.0	277	2	Q9F5T6	Q9F5T6 pseudomonas
22	1396	58.0	277	2	Q9F5S7	Q9F5S7 pseudomonas
23	1393	57.8	277	2	Q9F5T4	Q9F5T4 pseudomonas
24	1393	57.8	277	2	Q9F5T2	Q9F5T2 pseudomonas
25	1393	57.8	277	2	Q9F5T1	Q9F5T1 pseudomonas
26	1393	57.8	277	2	Q9F5S6	Q9F5S6 pseudomonas
27	1393	57.8	277	2	Q9F5T7	Q9F5T7 pseudomonas
28	1390	57.7	277	2	Q9F5T7	Q9F5T7 pseudomonas
29	1384	57.5	277	2	Q9F5S2	Q9F5S2 pseudomonas
30	1355.5	56.3	450	2	Q9ZHH3	Q9ZHH3 burkholderi
31	1139	47.3	226	2	Q8VUD6	Q8VUD6 ralstonia s
32	1134	47.1	214	2	Q8VR23	Q8VR23 pseudomonas
33	1134	47.1	214	2	Q8VR24	Q8VR24 pseudomonas
34	1127	46.8	214	2	Q8VR24	Q8VR24 pseudomonas
35	1123	46.6	214	2	Q8VR22	Q8VR22 pseudomonas
36	1113.5	46.2	455	2	Q93NA8	Q93NA8 burkholderi
37	1067	44.3	214	2	Q8VR25	Q8VR25 burkholderi
38	1045.5	43.4	437	2	Q9WXG8	Q9WXG8 alcaligenes
39	1040	43.2	208	2	Q9F6B6	Q9F6B6 marinobacte
40	1039	43.1	206	2	Q8VUM5	Q8VUM5 naphthalene
41	1038	43.1	214	2	Q8VR21	Q8VR21 burkholderi
42	1038	43.1	303	2	Q8VUD2	Q8VUD2 ralstonia s
43	1016	42.2	297	2	Q93M40	Q93M40 uncultured
44	1009	41.9	297	2	Q93M39	Q93M39 uncultured
45	1002	41.6	297	2	Q93M41	Q93M41 uncultured

## ALIGNMENTS

RESULT 1	Q938R9	PRELIMINARY:	PRT:	449 AA.
ID	Q938R9			
AC	Q938R9:			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Naphthalene dioxygenase.			
GN	NAHAC.			
OS	Pseudomonas fluorescens.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OX	Pseudomonas.			
NCBI_taxid=294;				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Min K.-H., Ji S.-H.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY048759; AAL07262.1; -.			
DR	InterPro; IPR001281; Rieseke.			
DR	InterPro; IPR001663; Ring_hydroxyl_A.			
DR	Pfam; PF00355; Rieseke; 1.			
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.			
DR	PROSITE; PS00570; Ring_HYDROXYL_ALPHA; UNKNOWN_1.			
KW	Dioxygenase.			
SO	SEQUENCE 449 AA; 49608 MW; 5EFDDE0282FE812 CRC64;			
Query Match	96.8%;	Score 2332;	DB 2;	Length 449;
Best local similarity	96.0%;	Pred. No. 1e-171;		
Matches 431;	Conservative 12;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1	MNNYKILVSESGLSQKHLIHGDELFQHELTFTFARMWLFTHDSLTPADGVYATMG	60	
DB	1	MNNYKILVSESGLSQKHLIHGDELFQHELTFTFARMWLFTHDSLTPADGVYATMG	60	
QY	61	IDEVYVSQNDGSTRAPFNVCRHKGKTLVSVAGNAKGFVCSYHGWGSGNGEIQSVPE	120	
DB	61	IDEVYVSQNDGSTRAPFNVCRHKGKTLVSVAGNAKGFVCSYHGWGSGNGEIQSVPE	120	
QY	121	KDLVGESLNKKCLGKEVARVESFPGFIYGCDFDEAPPLMDYLDAAVYLDPMKRKHSGL	180	

```

Db 121 KELGESLNKKCLCLEKAVRSEFHGFIYGCDEAPSLIMYLGAANYLEPIFKHSGL 180
QY 181 ELVGPCKVYVKAMKKAPENFVGDAYHVGWTHASSLSRSGSFSSLAGNALPPEGAGL 240
Db 181 ELVGPCKVYVKAMKKAPENFVGDAYHVGWTHASSLSRSGSFSSLAGNALPPEGAGL 240
QY 241 QMTSKYSGMGVLMWDGYSGVHSADLVPBELMAFGAKOERLKEIGDVRARIYRSHLNTCV 300
Db 241 QMTSKYSGMGVLMWDGYSGVHSADLVPBELMAFGAKOERLKEIGDVRARIYRSHLNTCV 300
QY 301 FPNNSMILTCSCGVFKVWNPIDANTTEWVTYAIYKDMPEDLKRLADSVQRTLPAGFWE 360
Db 301 FPNNSVLTCSCGVFKVWNPIDANTTEWVTYAIYKDMPEDLKRLADSVQRTLPAGFWE 360
QY 361 DDNDNMETASONGKKYOSRSDLSNLGFGEDYVDGDAYPGVVGKSAIGETSYRGYRAY 420
Db 361 DDNDNMETASONGKKYOSRSDLSNLGFGEDYVDGDAYPGVVGKSAIGETSYRGYRAY 420
QY 421 QAHVSSSNMAEFEDASSFWHTELTKTTDR 449
Db 421 QAHVSSSNMAEFEDASSFWHTELTKTTDR 449

```

```

RESULT 2
Q52142 PRELIMINARY; PRT; 449 AA.
AC 052142; 008194;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Iron-sulfur protein large subunit.
GN PAHAC.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=903;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA Takizawa N.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA MEDLINE=94209249; PubMed=8157615;
RA TAKIZAWA N., Kaide N., Torioge S., Moritani T., Sawada T., Satoh S.,
RA Kiyohara H.;
RT "Identification and characterization of genes encoding polycyclic
RT aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon
RT dihydrodiol dehydrogenase in Pseudomonas putida OUS82.";
RL J. Bacteriol. 176:2444-2449(1994).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN=OUS82;
RA Noboru T., Toshiya I., Takashi S., Kazuhiro Y., Masamichi K.,
RA Yue-Wu W., Masao F., Hozoh K.;
RT "The molecular analysis of NAH7-type cluster located on the
RT chromosomes of Pseudomonas aeruginosa Pak1 and Pseudomonas putida
RT OUS82.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN 14
RP EMBL; AB004059; BAA20391.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 449 AA; 49361 MW; 80020F54AAB11EBA CRC64;
Query Match 95.2%; Score 2292; DB 2; Length 449;
Best Local Similarity 94.4%; Pred. No. 1,3e-168;

```

```

Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;
QY 1 MNYNKILVSESGLSOKHLIHGDEELFOHELKTFARNMWLFTHDSLIPAPGDYVAKWG 60
Db 1 MNYNKILVSESGLSOKHLIHGDEELFOHELKTFARNMWLFTHDSLIPAPGDYVAKWG 60
QY 61 IDEVIVSRQNDGSIIRAFNLNVCRRHAKTLVSEAGNAKGFVCSYHGMOFGSNGELQSVPE 120
Db 61 IDEVIVSRQNDGSIIRAFNLNVCRRHAKTLVSEAGNAKGFVCSYHGMOFGSNGELQSVPE 120
QY 121 KDLYGESLNKKCLCLEKAVRSEFHGFIYGCDEAPSLIMYLGAANYLEPIFKHSGL 180
Db 121 KDLYGESLNKKCLCLEKAVRSEFHGFIYGCDEAPSLIMYLGAANYLEPIFKHSGL 180
QY 181 ELVGPCKVYVKAMKKAPENFVGDAYHVGWTHASSLSRSGSFSSLAGNALPPEGAGL 240
Db 181 ELVGPCKVYVKAMKKAPENFVGDAYHVGWTHASSLSRSGSFSSLAGNALPPEGAGL 240
QY 241 QMTSKYSGMGVLMWDGYSGVHSADLVPBELMAFGAKOERLKEIGDVRARIYRSHLNTCV 300
Db 241 QMTSKYSGMGVLMWDGYSGVHSADLVPBELMAFGAKOERLKEIGDVRARIYRSHLNTCV 300
QY 301 FPNNSMILTCSCGVFKVWNPIDANTTEWVTYAIYKDMPEDLKRLADSVQRTLPAGFWE 360
Db 301 FPNNSVLTCSCGVFKVWNPIDANTTEWVTYAIYKDMPEDLKRLADSVQRTLPAGFWE 360
QY 361 DDNDNMETASONGKKYOSRSDLSNLGFGEDYVDGDAYPGVVGKSAIGETSYRGYRAY 420
Db 361 DDNDNMETASONGKKYOSRSDLSNLGFGEDYVDGDAYPGVVGKSAIGETSYRGYRAY 420
QY 421 QAHVSSSNMAEFEDASSFWHTELTKTTDR 449
Db 421 QAHVSSSNMAEFEDASSFWHTELTKTTDR 449

```

```

RESULT 3
Q92173 PRELIMINARY; PRT; 449 AA.
AC 092173;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase Fe-S large subunit.
GN NAHAC.
OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=316;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=AN10;
RX MEDLINE=99365311; PubMed=10433976;
RA Bosch R., Garcia-Valdes E., Moore E.R.B.;
RT "Genetic characterization and evolutionary implications of a
RT chromosomally encoded naphthalene-degradation upper pathway from
RT Pseudomonas stutzeri AN10.";
RL Gene 236:149-157(1999).
DR EMBL; AF039533; AB02136.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
SQ SEQUENCE 449 AA; 49802 MW; BA510FD1B6F6E63 CRC64;
Query Match 91.7%; Score 2208; DB 2; Length 449;
Best Local Similarity 89.5%; Pred. No. 3.8e-162;
Matches 402; Conservative 25; Mismatches 22; Indels 0; Gaps 0;
QY 1 MNYNKILVSESGLSOKHLIHGDEELFOHELKTFARNMWLFTHDSLIPAPGDYVAKWG 60

```



```
Db 1 MNVKNKTLVTESGLTQKHLIHGDEELFOHELETFIPARNMLFLTHDSLIPSPDYVTAKMG 60
QY 61 IDEVIVSRONDGSIKRAFLNVCRRHCKTLIVSAGNAKGVCSYHGMGRSNGELOSVPFE 120
Db 61 VDEVIVSRONDGSIKRAFLNVCRRHCKTLIVHAEGNAKGVCSYHGMGRSNGELOSVPFE 120
QY 121 KOLYGESLNKKCLGLKEVARVESFHGFIYGCDFDQAPPLMDYLGDAAMYLEBPMFKHSGGL 180
Db 121 KELYEPGKLVIIKANKKAPAEFNVGDAYHYGWTHTASSLSRSGESIFSSLAGNMALEPPEGAGL 180
QY 181 ELVGPFGKVIYIANKKAPAEFNVGDAYHYGWTHTASSLSRSGESIFSSLAGNMALEPPEGAGL 240
Db 181 ELVGPFGKVIYIANKKAPAEFNVGDAYHYGWTHTASSLSRSGESIFSSLAGNMALEPPEGAGL 240
QY 241 OMTSKYSGMGVLMGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYRSHLNCV 300
Db 241 OMTSKYSGMGVLMGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYRSHLNCV 300
QY 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVQRTLGPAFGFMS 360
Db 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVQRTLGPAFGFMS 360
QY 361 DDNDNMETASONGKRYQSRSDDLNLGFGEDYGDVAYPGVYKSAIGETSYRGFYRAY 420
Db 361 DDNDNMETASONGKRYQSRSDDLNLGFGEDYGDVAYPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449
Db 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449

RESULT 4
QY 09L5D2 PRELIMINARY: PRT: 447 AA.
AC 09L5D2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Polyaromatic hydrocarbon dioxygenase large subunit.
GN PAHAC.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxId=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H;
RX MEDLINE=21306905; PubMed=11414329;
RA Moser R., Stahl U.;
RT "Insights into the genetic diversity of initial dioxygenases from PAH-
  degrading bacteria.";
RL Appl. Microbiol. Biotechnol. 55:609-618(2001).
DR EMBL; AF252550; AAF72976.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR PRINTS; PR00090; Ring_hydroxyl_A; 1.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
SQ SEQUENCE 447 AA; 49548 MW; 2950EF36123A6F21 CRC64;

Query Match 91.1%; Score 2193; DB 2; Length 447;
Best Local Similarity 88.9%; Pred. No. 5.5e-16;
Matches 399; Conservative 27; Mismatches 21; Indels 2; Gaps 1;
```

```
Db 59 VDEVIVSRONDGSIKRAFLNVCRRHCKTLIVHAEGNAKGVCSYHGMGRSNGELOSVPFE 118
QY 121 KOLYGESLNKKCLGLKEVARVESFHGFIYGCDFDQAPPLMDYLGDAAMYLEBPMFKHSGGL 180
Db 119 KELYEPGKLVIIKANKKAPAEFNVGDAYHYGWTHTASSLSRSGESIFSSLAGNMALEPPEGAGL 178
QY 181 ELVGPFGKVIYIANKKAPAEFNVGDAYHYGWTHTASSLSRSGESIFSSLAGNMALEPPEGAGL 240
Db 179 ELVGPFGKVIYIANKKAPAEFNVGDAYHYGWTHTASSLSRSGESIFSSLAGNMALEPPEGAGL 238
QY 241 OMTSKYSGMGVLMGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYRSHLNCV 300
Db 239 OMTSKYSGMGVLMGYSGVHSADLVPELMAFGAKOEKRLKEIGDVARIRYRSHLNCV 298
QY 301 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVQRTLGPAFGFMS 360
Db 299 FPNNSMLTCSGVFKYWNPIDANTTEVWTYAIYEKMPEDLKRRLADSVQRTLGPAFGFMS 358
QY 361 DDNDNMETASONGKRYQSRSDDLNLGFGEDYGDVAYPGVYKSAIGETSYRGFYRAY 420
Db 359 DDNDNMETASONGKRYQSRSDDLNLGFGEDYGDVAYPGVYKSAIGETSYRGFYRAY 418
QY 421 QAHVSSNMWAEFEHASTWHTELTKTTDR 449
Db 419 QAHVSSNMWAEFEHASTWHTELTKTTDR 447

RESULT 5
QY 052382 PRELIMINARY: PRT: 447 AA.
AC 052382;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene 1,2 dioxygenase large oxygenase component.
GN NAGAC.
OS Ralstonia sp. U2.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxId=70356;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U2;
RX MEDLINE=98233751; PubMed=9573207;
RA Fuenmayor S.L., Wild M., Boyes A.L., Williams P.A.;
RT "A gene cluster encoding steps in conversion of naphthalene to
  gentisate in Pseudomonas sp. strain U2.";
RL J. Bacteriol. 180:2522-2530(1998).
RP SEQUENCE FROM N.A.
RC STRAIN=U2;
RX MEDLINE=20576173; PubMed=11133965;
RA Zhou N.Y., Fuenmayor S.L., Williams P.A.;
RT "nag genes of Ralstonia (Formerly pseudomonas) sp. Strain U2 encoding
  enzymes for gentisate catabolism.";
RL J. Bacteriol. 183:700-708(2001).
DR EMBL; AF036940; AAD12610.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR PRINTS; PR00090; Ring_hydroxyl_A; 1.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase; Plasmid.
SQ SEQUENCE 447 AA; 49570 MW; 4553AAF4410ED0 CRC64;

Query Match 90.9%; Score 2189; DB 2; Length 447;
Best Local Similarity 89.1%; Pred. No. 1.1e-16;
Matches 400; Conservative 25; Mismatches 22; Indels 2; Gaps 1;
```

```

Db      1 M1TEN - LVSEGLQ1KHFLHGDKELEFHEKHT1FARNWLE1LTHDSL1PBGDTV1AKKG 58
Oy      61 IDEIVSVRONDSIRAF1LNVCRHNRKQITV1SVACNAKGFVCSYHGWFSGNGELQ5YPE 120
Db      59 VDEV1SVRONDSVAF1LNVCRHNRKQITV1VAHAEAGNAKGFVCSYHGWFSGNGELQ5YPE 118
Oy      121 KDLYSESLNKKCLGLEVARVSEFHF1YCGFDDAPPLDYLDGDAWYLEPMFKHSGGL 180
Db      119 KELYGDT1KKKCLGLEKVEPRIESFHF1YCGFDDAPPLDYLDGDAWYLEPMFKHSGGL 178
Oy      181 ELYVPGPKVY1KANKKAPAE1NVCGDAHVGCTHA1SRGSEJTFSS1AGNAALPPBAGL 240
Db      179 ELYVPGPKVY1KANKKAPAE1NVCGDAHVGCTHA1SRGSEJTFPLAGNAMPPEBAGL 238
Oy      241 QMTSVYSGMGV1MDGYSGVSHAD1VPE1LMAFGAKOBLKE1GVDAR1YSH1NCVY 300
Db      239 QMTSVYSGMGV1MDGYSGVSHAD1VPE1LMAFGAKOBLKE1GVDAR1YSH1NCVY 298
Oy      301 FPNNSMLTCSGVYFKWPN1DANTTEVWY1A1VEKDMPEDLKRL1ADSVQTL1PAGFWE 360
Db      299 FPNNSI1LTCSGVYFKWPN1IDENTTEVWY1A1VEKDMPEDLKRL1ADSVQTL1PAGFWE 358
Oy      361 DDNDNMETASONGKRRYQSDSL1NLGFGEDVYGDAY1PBGVYKSA1GETSVRGFYRAY 420
Db      359 DDNDNMETESQAKKRYQSSNDL1ANLGFGEDVYGDCEYGPVYAKSA1GETSVRGFYRAY 418
Oy      421 QAHVSSWMAEPHASS1TWHTEL1KTTDR 449
Db      419 QAH1SSWMAEPHASS1RWHTEL1KTTDR 447

```

```

Db      1 M5YQ~--LVSEAGLQKHLITYGDKELFQHEIKTJTFARNWLELTHDSLIPSGDVIYKAMG 58
Qy      61 IDEIVYSRONDISIRAPLINVCRHCKTLTVSEAGNAKGFVCSYGMGFGSNGELQ5VPE 120
Db      59 VDEIVYSRONDSVAFELNVCRHCKTITVDADAGNAKGFVCGYGMKGGSNGELQ5VPE 118
Qy      121 KDLYGESLNKCKLGLKEVARVESFGEFTYGGCFDQAPLMDYLDLDAAMYLEPMKHSGL 180
Db      119 KELYGDATIKKCKLGLKEVPRIESFGEFTYGGCFDQAPLIDYLDDAAMYLEPMKHSGL 178
Qy      181 ELVGPBGKVIYKANKKAPENFVGDAVYGMTHASSLSRGSIFSSLAGNALPPEGAGL 240
Db      179 ELVGPBGKVVYKANKPLAENFVGDAVYHIGMTHASILRAGSIFAPLAGNAPEPEGAGL 238
Qy      241 QMTSKYSGSMGLVMGCSGVHSADLVPELMLPFGAKOKORLKEIGDVRARIYRHLNCTV 300
Db      239 QMTTKYSGIGLVMAYGCSIGSADVPMPMAFGAKOKELKEIGDVRARIYR50LNGTV 298
Qy      301 FPNNSMLTFCSGYFKVWNPIDANTTEVWYIAYVEKDMPEDLKRLRADS5ORTLPAGFWES 360
Db      299 FPNNSFLTCGYFKVFNPIDENTTEVWYIAYVEKDMPEDLKRLRADA5QSVGAGWES 358
Qy      361 DDNDMMETA5QNGKKYQSRSDLSNLGFGEDVYGDAVPGCVKSAIGETSYGFFRAY 420
Db      359 DDNDMMGLT5QNAKKYQSSNDLADLGFQGDVYGEDEYCVGVKSAISESTYGFYRAY 418
Qy      421 QAHVSS5MAAEFEHAA5STWHTELTKTDR 449
Db      419 QAHSS5MAAEFE5SRMHTEELTKTDR 447

```

RESULT	6			
Q8VUD4				
Q8VUD4	PRELIMINARY;	PRT:	447	AA.
AC	Q8VUD4;			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	DntC.			
OS	Burkholderia cepacia (Pseudomonas cepacia).			
OC	Bacteria; Proteobacteria; beta subdivision; Burkholderia group;			
CC	Burkholderia.			
OX	NCBI_TaxID=292;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=R34;			
RX	MEDLINE=20254695; PubMed=10795678;			
RA	Johnson G.R., Jain R.K., Spain J.C.;			
RT	"Properties of the trihydroxytoluene oxygenase from Burkholderia			
RT	cepacia R34: an extradiol dioxygenase from the 2,4-dinitrotoluene			
RT	pathway.";			
RL	Arch. Microbiol. 173:86-90(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=R34;			
RA	Johnson G.R., Jain R.K., Spain J.C.;			
RT	"Origins of the 2,4-dinitrotoluene pathway.";			
RL	Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL, AF169302; AAL50021.1; .			
DR	InterPro; IPR001281; Rlseske.			
DR	InterPro; IPR001663; Ring_hydroxyl_A.			
DR	Pfam; PF00355; Rlseske; 1.			
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.			
DR	PRINTS; PR00090; RINGIOXGNASE.			
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.			
DR	PROSITE; 4447 AA; 49468 MW; 2A5BF558320275C CRC64;			
QO	SEQUENCE			

RESULT	7
ID	P95564
AC	P95564
DT	01-MAY-1997 (TREMBLrel. 03, Created)
DT	01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE	IsipAlpha 2NT.
GN	NRDAG.
OS	<i>Pseudomonas</i> sp.
OC	Bacteria; Proteobacteria.
OX	NCBI_TaxID=306;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JS42;
RA	MEDLINE=97128768; PubMed=8973308;
RT	Parales J.V., Kumar A., Parales R.E., Gibson D.T.;
RT	"Cloning and sequencing of the genes encoding 2-nitrotoluene
RT	dioxygenase from <i>Pseudomonas</i> sp. JS42."
RL	Gene 181:57-61(1996).
DR	EMBL: U49504; AAB40383.1; .
DR	HSSP: P23094; INDO.
DR	InterPro: IPR001281; Rieske.
DR	InterPro: IPR001663; Ring_hydroxyl_A.
DR	Pfam: PF00355; Rieske; 1.
DR	Pfam: PF00848; Ring_hydroxyl_A; 1.
DR	PRINTS: PR00090; RINGDIOXGNASE.
DR	PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
SO	SEQUENCE 447 AA; 49485 MW; 1CB0E223E528E3BD CRC64;
Query March	85.3%; Score 2054; DB 2; Length 447;
Best Local Similarity	83.7%; Pred. NO. 2.9e-150;
Matches 376; Conservative	35; Mismatches 36; Indels 2; Gaps 1;
1	MYNNKILIVEESLSQKHILHGDEELFOHEKKTIFARNMILFTLTHDSLIPAPGDYVAKMG 60

Query Match	86.0%;	Score 2072;	DB 2;	length 447;
Best Local Similarity	83.5%;	Pred. NO.1.2e-151;		
Matches 375;	Conservative 39;	Mismatches 33;	Indels 2;	Gaps 1;
0y	1	MNYNNKILVSESGLSOKHLIHGDEELFQHELKTIIFARNMLFTVHDSLLIPAPGDYVTAKNG	60	

Db 1 MSYQ--LVSEAGLTQKLLIHGDKELFQHELTIFARNWLEFLTHDSSLPSGDIYVAKMG 58  
Qy 61 IDEVIVSRQNDGSIAPFLPNCVHRHKGTLTVSEAGNAKPFVCSYHGWFSGNGEIOSPPE 120  
Db 59 VDEVIVSRQNDGSIAPFLPNCVHRHKGTLTVHTEAGNAGPFVCSYHGWFSGNGEIOSPPE 118

RESULT	8			
Q8RTL4				
ID	Q8RTL4	PRELIMINARY;	PRT;	447 AA.
AC	Q8RTL4;			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	Oxygenase-alpha NBD0.			
GN	NB2AC.			
OS	Comamonas sp. JS765.			
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas			
OX	NCBI_TaxID=58226;			
	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JS765;			
RX	MEDLINE=21681044; PubMed=11823201;			
RA	Lesner D.J., Johnson G.R., Parales R.E., Spain J.C., Gibson D.T.;			
RT	"Molecular Characterization and Substrate Specificity of Nitrobenzene			
RT	Dioxygenase from Comamonas sp. Strain JS765.";			
RL	Appl. Environ. Microbiol. 68:634-641(2002).			
SD	EMBL: AF379638; AAL76202.1; -			
SD	SEQUENCE 447 AA; 49556 MW; 4CA961F99A37DF3A CRC64;			

[illegible]

RESULT 9			
ID	Q45695	PRELIMINARY;	PRT; 451 AA.
AC	Q45695;		
DT	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	ISP-alpha.		
GN	DNTAC.		
OS	Burkholderia sp. (strain RASC).		
CC	Bacteria; Proteobacteriia; beta subdivision; Burkholderia group;		
OC	Burkholderia.		
OX	NCBI_TaxID=69003;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=DNT;		
RX	MEDLINE=96326338; PubMed=8759857;		
RA	Suen W.C., Haigler B.E., Spain J.C.;		
RT	"2,4-Dinitrotoluene dioxygenase from Burkholderia sp. strain DNT:		
RT	similarity to naphthalene dioxygenase.";		
RL	J. Bacteriol. 178:4926-4934(1996).		
DR	EMBL: U62430; AAB09766.1; -.		
DR	HSSP; P23094; INDO.		
DR	InterPro; IPR001281; Rieske.		
DR	InterPro; IPR001663; Ring_hydroxyl_A.		
DR	Pfam; PF00355; Rieske; 1.		
DR	Pfam; PF00848; Ring_hydroxyl_A; 1.		
DR	PRINTS; PR00090; RNCDIOXGNASE.		
DR	PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.		
SO	SEQUENCE 451 AA: 49827 MW: 726796C29C9F9A10 CRC64;		

[illegible]

```
||||| :|||||:||||| ||||| ||:|||||
Db 364 DDNDNM-VLSGNARKYSSNSDLADLFGRKDYGDCEYGPVSKSAPSETHNRGFYRAY 422
QY 421 QAHSSSNMAFEFFHASSSTWHTLTFTTDR 449
||:|||||:| |||||||
Db 423 QAHSSSNMAFEFFHASSSTWHTLTFTTDR 451

RESULT 10
O9ETK2 PRELIMINARY; PRT; 277 AA.
ID O9ETK2:
AC O9ETK2:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=2JIDNH, PRIMNL AND 3JA2NH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF306438; AAC25698.1; -.
DR EMBL: AF306432; AAC25692.1; -.
DR EMBL: AF306436; AAC25696.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30094 MW; BFCF9B5A379546DB CRC64;

Query Match 61.9%; Score 1490; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 4.9e-107;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRONDGSIKRAFLVNCRRHCKTIVSVEAGNAKGFVCSYHGMGFGSGNGLQ 115
|||||
Db 1 TAKMGIDEVIVSRONDGSIKRAFLVNCRRHCKTIVSVEAGNAKGFVCSYHGMGFGSGNGLQ 60
QY 116 SVPEFKDLVGSLLKKKCLGKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAWYLEPMFK 175
|||||
Db 61 SVPEFKDLVGSLLKKKCLGKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAWYLEPMFK 120
QY 176 HSGGLELVGPPGKVIVITANKKAPAEVFGDAYHVGWTHASSLSRGSSTFSSLAGNAALPP 235
|||||
Db 121 HSGGLELVGPPGKVIVITANKKAPAEVFGDAYHVGWTHASSLSRGSSTFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPFLMAGKAGQERLKEIGDVARIRYRSH 295
|||||
Db 181 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPFLMAGKAGQERLKEIGDVARIRYRSH 240
QY 296 LNCVFPNNSMLTCSGVFKVWNPIDANTTEWYTAIV 332
|||||
Db 241 LNCVFPNNSMLTCSGVFKVWNPIDANTTEWYTAIV 277

RESULT 11
O9F5S4 PRELIMINARY; PRT; 277 AA.
ID O9F5S4:
AC O9F5S4:
DT 01-MAR-2001 (Tremblrel. 16, Created)
```

```
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=5IITASAL;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF306440; AAC25700.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1
FT NON_TER 277
SQ SEQUENCE 277 AA; 30094 MW; FADFDA5D37E241AB CRC64;

Query Match 61.8%; Score 1488; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 7e-107;
Matches 276; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIVSRONDGSIKRAFLVNCRRHCKTIVSVEAGNAKGFVCSYHGMGFGSGNGLQ 115
|||||
Db 1 TAKMGIDEVIVSRONDGSIKRAFLVNCRRHCKTIVSVEAGNAKGFVCSYHGMGFGSGNGLQ 60
QY 116 SVPEFKDLVGSLLKKKCLGKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAWYLEPMFK 175
|||||
Db 61 SVPEFKDLVGSLLKKKCLGKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAWYLEPMFK 120
QY 176 HSGGLELVGPPGKVIVITANKKAPAEVFGDAYHVGWTHASSLSRGSSTFSSLAGNAALPP 235
|||||
Db 121 HSGGLELVGPPGKVIVITANKKAPAEVFGDAYHVGWTHASSLSRGSSTFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPFLMAGKAGQERLKEIGDVARIRYRSH 295
|||||
Db 181 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPFLMAGKAGQERLKEIGDVARIRYRSH 240
QY 296 LNCVFPNNSMLTCSGVFKVWNPIDANTTEWYTAIV 332
|||||
Db 241 LNCVFPNNSMLTCSGVFKVWNPIDANTTEWYTAIV 277

RESULT 12
O9F5T3 PRELIMINARY; PRT; 277 AA.
ID O9F5T3:
AC O9F5T3:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC.
OS Pseudomonas sp. 19IIDNH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139738;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=19IIDNH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL: AF306426; AAG25686.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Dioxigenase.
KW NON_TER 1
FT NON_TER 1
SQ SEQUENCE 277 AA; 30064 MW; DB47868EB6DD525A CRC64;

Query Match
Best Local Similarity 61.7%; Score 1486; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDDEVIVSRQNDGSIKRAFLNVCRRHKGKTLVSVYDAGNAKGFVCSYHGWGFSGNGELQ 115
DB 1 TAKMGIDDEVIVSRQNDGSIKRAFLNVCRRHKGKTLVSVYDAGNAKGFVCSYHGWGFSGNGELQ 60

QY 116 SVPEFKDLVGEISLNKKCLGKEVARVESFHGFTYGCEDQAPPLMDYLGDAAMYLEPMFK 175
DB 61 GVPEFKDLVGEISLNKKCLGKEVARVESFHGFTYGCEDQAPPLMDYLGDAAMYLEPMFK 120

QY 176 HSGGLELVGPPGKVVYIKAMMKAPAEENVGDYHVGTWTHASSLSRSGSIFSSLAGNALPP 235
DB 121 HSGGLELVGPPGKVVYIKAMMKAPAEENVGDYHVGTWTHASSLSRSGSIFSSLAGNALPP 180

QY 236 EGAGLQMTSKYSGGMGLVDGYSGVSHADLVPELMAFGAKQERLNKEIGDVARIRYRSH 295
DB 181 EGAGLQMTSKYSGGMGLVDGYSGVSHADLVPELMAFGAKQERLNKEIGDVARIRYRSH 240

QY 296 LNCVFPPNNSMLTCGSGVFKVWNPIDANTTEVWTYAIY 332
DB 241 LNCVFPPNNSMLTCGSGVFKVWNPIDANTTEVWTYAIY 277

RESULT 13
Q9F5S8 PRELIMINARY; PRT; 277 AA.
AC Q9F5S8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Naphthalene dioxigenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. PR3MN2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR3MN2;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxigenase genes (nahAc) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF306434; AAG25694.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Dioxigenase.
KW NON_TER 1
FT NON_TER 1
SQ SEQUENCE 277 AA; 30064 MW; 6BFF83D32F04CFC3 CRC64;

Query Match
Best Local Similarity 61.6%; Score 1484; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDDEVIVSRQNDGSIKRAFLNVCRRHKGKTLVSVYDAGNAKGFVCSYHGWGFSGNGELQ 115
DB 1 TAKMGIDDEVIVSRQNDGSIKRAFLNVCRRHKGKTLVSVYDAGNAKGFVCSYHGWGFSGNGELQ 60

QY 116 SVPEFKDLVGEISLNKKCLGKEVARVESFHGFTYGCEDQAPPLMDYLGDAAMYLEPMFK 175
DB 61 SVPEFKDLVGEISLNKKCLGKEVARVESFHGFTYGCEDQAPPLMDYLGDAAMYLEPMFK 120

QY 176 HSGGLELVGPPGKVVYIKAMMKAPAEENVGDYHVGTWTHASSLSRSGSIFSSLAGNALPP 235
DB 121 HSGGLELVGPPGKVVYIKAMMKAPAEENVGDYHVGTWTHASSLSRSGSIFSSLAGNALPP 180

QY 236 EGAGLQMTSKYSGGMGLVDGYSGVSHADLVPELMAFGAKQERLNKEIGDVARIRYRSH 295
DB 181 EGAGLQMTSKYSGGMGLVDGYSGVSHADLVPELMAFGAKQERLNKEIGDVARIRYRSH 240

QY 296 LNCVFPPNNSMLTCGSGVFKVWNPIDANTTEVWTYAIY 332
DB 241 LNCVFPPNNSMLTCGSGVFKVWNPIDANTTEVWTYAIY 277

RESULT 14
Q9F5S9 PRELIMINARY; PRT; 277 AA.
AC Q9F5S9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Naphthalene dioxigenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 8IDINH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8IDINH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxigenase genes (nahAc) in
Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF306433; AAG25693.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Dioxigenase.
KW NON_TER 1
FT NON_TER 1
SQ SEQUENCE 277 AA; 29995 MW; AADA9B5A378F6DB CRC64;

Query Match
Best Local Similarity 61.6%; Score 1483; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDDEVIVSRQNDGSIKRAFLNVCRRHKGKTLVSVYDAGNAKGFVCSYHGWGFSGNGELQ 115
DB 1 TAKMGIDDEVIVSRQNDGSIKRAFLNVCRRHKGKTLVSVYDAGNAKGFVCSYHGWGFSGNGELQ 60

QY 116 SVPEFKDLVGEISLNKKCLGKEVARVESFHGFTYGCEDQAPPLMDYLGDAAMYLEPMFK 175
DB 61 SVPEFKDLVGEISLNKKCLGKEVARVESFHGFTYGCEDQAPPLMDYLGDAAMYLEPMFK 120

QY 176 HSGGLELVGPPGKVVYIKAMMKAPAEENVGDYHVGTWTHASSLSRSGSIFSSLAGNALPP 235
DB 121 HSGGLELVGPPGKVVYIKAMMKAPAEENVGDYHVGTWTHASSLSRSGSIFSSLAGNALPP 180

QY 236 EGAGLQMTSKYSGGMGLVDGYSGVSHADLVPELMAFGAKQERLNKEIGDVARIRYRSH 295
```

Db 181 EGAGLQMTSKYSGGVLMDGSGVSHADLVPBELMAFGAKOGLNKKEIGDVRARIYRSH 240  
QY 296 LNCITVFPNNNSMLTCSGVFKVWNPIDANTTEVWTAIV 332  
Db 241 LNCITVFPNNNSMLTCSGVFKVWNPIDANTTEVWTAIV 277

## RESULT 15

Q9F5S3 PRELIMINARY: PRT: 277 AA.  
AC Q9F5S3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 1 (Fragment).  
GN NAHAC1.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=511AH;  
RA Ferrero M.A., Jalucat J., Bosch R.;  
RT "Coexistence of two naphthalene dioxygenase genes (nahAc) in  
Pseudomonas strains from West Mediterranean Sea."  
RL Submitted (SFP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF306441; AAC25701.1; .  
DR HSSP; P23094; INDO.  
DR InterPro; IPR001281; Rieske.  
DR InterPro; IPR001663; Ring\_hydroxyl\_A.  
DR Pfam; PF00355; Rieske; 1.  
DR Pfam; PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS; PR00090; RINGDIOXGNASE.  
DR PROSITE; PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
FT NON\_TER 1 1  
FT 277 277  
SQ SEQUENCE 277 AA; 30062 MW; 82C7BD865A92D0AB CRC64;

Query Match 60.3%; Score 1451; DB 2; Length 277;  
Best Local Similarity 96.8%; Pred. No. 5e-104;  
Matches 268; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 56 TAKKGIDEVITYSRNDGSIKRAFLNVCRRHRTLYVEAGNAKGFVCSYHGWGFGSNGELQ 115  
Db 1 TAKKGIDEVITYSRNDGSIKRAFLNVCRRHRTLYVEAGNAKGFVCSYHGWGFGSNGELQ 60  
QY 116 SVPEKDLYGESLKKKCLGLKEVARVESFHGFIYGCFOEAPPLMDYLGDAAWYLEPMFK 175  
Db 61 SVPEKELYGESLKKKCLGLKEVARVESFHGFIYGCFOEAPSLMDYLGDAAWYLEPIFK 120  
QY 176 HSGGLELVGPPGKVIYKANMKAPAEENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPP 235  
Db 121 HSGGLELVGPPGKVIYKANMKAPAEENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPP 180  
QY 236 EGAGLQMTSKYSGGVLMDGSGVSHADLVPBELMAFGAKOERLNKEIGDVRARIYRSH 295  
Db 181 EGAGLQMTSKYSGGVLMDGSGVSHADLVPBELMAFGAKOERLNKEIGDVRARIYRSH 240  
QY 296 LNCITVFPNNNSMLTCSGVFKVWNPIDANTTEVWTAIV 332  
Db 241 LNCITVFPNNNSMLTCSGVFKVWNPIDANTTEVWTAIV 277

Search completed: January 27, 2003, 09:02:23  
Job time : 26.8333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:47:32 : Search time 7.16667 Seconds  
(without alignments)  
2538.540 Million cell updates/sec

Title: US-09-843-250-35

Perfect score: 2408

Sequence: 1 MNVNNKILVSESGLSQKHLI.....AEFEHASTWHTELTKTTDR 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2404	99.8	449	1	NDOB_PSEPU
2	2318	96.3	449	1	NDOB_PSEPL
3	2219	92.2	449	1	NDOB_PSEAB
4	737	30.6	453	1	HCAE_ECOLI
5	652.5	27.1	450	1	BED1_PSEPU
6	650.5	27.0	450	1	TOD1_PSEPU
7	617.5	25.6	458	1	BPHA_BURCE
8	617	25.6	457	1	BPHA_COMTE
9	617	25.6	458	1	BPHA_PSEPS
10	588.5	24.4	448	1	BNZA_PSEPU
11	587	24.4	458	1	BNZA_PSESI
12	410.5	17.0	454	1	XYLX_PSEPU
13	399	16.6	461	1	BEWA_AICCA
14	251	10.4	374	1	YEAW_ECOLI
15	239.5	9.9	442	1	CHMO_AMATR
16	233	9.7	438	1	CHMO_ATRHO
17	224	9.3	445	1	CHMO_BEYU
18	223	9.3	439	1	CHMO_SPIOL
19	216	9.0	422	1	CHMO_ARATH
20	163	6.8	439	1	PH13_PSEPU
21	118.5	4.9	354	1	VANA_PSESP
22	110	4.6	329	1	VANA_PSESE9
23	110	4.6	432	1	CRAA_COMTE
24	109	4.5	1276	1	PMR6_CHLDP
25	103.5	4.3	1926	1	LPH_RABIT
26	102.5	4.3	468	1	PEP1_PICPA
27	102	4.2	409	1	POBA_PSEPS
28	101.5	4.2	3255	1	POLG_LMVE
29	99.5	4.1	543	1	7UP1_DROME
30	99.5	4.1	746	1	7UP2_DROME
31	98.5	4.1	1411	1	Y297_HUMAN
32	97.5	4.0	3255	1	POLG_LMVO
33	94.5	3.9	331	1	LDHA_RHIDE

34	94.5	3.9	350	1	UNRI_HUMAN
35	94.5	3.9	452	1	P26_YEAST
36	94.5	3.9	518	1	ATPA_ENTHR
37	94	3.9	847	1	ORP8_HUMAN
38	94	3.9	1693	1	POLN_HEVY
39	93	3.9	420	1	DHE3_PYRO
40	93	3.9	917	1	SVI_SPAU
41	92.5	3.8	475	1	SYC_AERPE
42	92.5	3.8	774	1	LOI2_HUMAN
43	92	3.8	643	1	SGT1_ARATH
44	92	3.8	985	1	AGLU_ASPOR
45	92	3.8	1693	1	POLN_HEVBU

## ALIGNMENTS

RESULT 1

ID	NDOB_PSEPU	STANDARD:	PRT:	449 AA.
AC	P23094: OS2124: O33461: O07830:			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene 1,2-dioxygenase ISP alpha)			
DE	NDOB OR NAHAC OR DOXB OR NAHA3 OR NDCC2.			
OS	Pseudomonas putida, and			
OS	Pseudomonas sp. (strain C18).			
OC	Plasmid pPTG1, Plasmid NAH7, and Plasmid NPL1.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_Taxid=303, 306;			
RM	[1]			
RM	SEQUENCE FROM N.A.			
RC	SPECIES=P. putida; STRAIN=NCIB 9816;			
RC	MEDLINE=89211973; PubMed=324348;			
RA	Kurkela S., Leivaslahti H., Palva E.T., Teeri T.H.;			
RT	"Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of Pseudomonas putida strain NCIB9816."			
RT	Gene 73:355-362(1988).			
RM	[2]			
RM	SEQUENCE FROM N.A.			
RC	SPECIES=P. putida; STRAIN=NCIB 9816;			
RC	Parales J.V., Parales R.E., Kumar A., Gibson D.T.;			
RA	Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.			
RM	[3]			
RM	SEQUENCE FROM N.A.			
RC	STRAIN=C18;			
RC	MEDLINE=94042852; PubMed=8226631;			
RA	Denome S.A., Stanley D.C., Olson E.S., Young K.D.;			
RT	"Metabolism of dibenzothiophene and naphthalene in Pseudomonas strains: complete DNA sequence of an upper naphthalene catabolic pathway."			
RT	J. Bacteriol. 175:6890-6901(1993).			
RM	[4]			
RM	SEQUENCE FROM N.A.			
RC	SPECIES=P. putida; STRAIN=G7 / ATCC 17485; PLASMID=NAH7;			
RC	MEDLINE=9325277; PubMed=8486285;			
RA	Simon M.J., Ossling T.D., Saunders R., Ensley B.D., Suggs S.;			
RA	Harcourt A.A., Shen W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;			
RT	"Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida strains G7 and NCIB 9816-4."			
RT	Gene 127:31-37(1993).			
RM	[5]			
RM	SEQUENCE FROM N.A.			
RC	SPECIES=P. putida; STRAIN=BS202; PLASMID=NPL1;			
RA	Bezdorodnikov S.G., Boronin A.M., Tiedje J.M.;			
RT	"Nucleotide sequences of genes encoding an upper pathway of naphthalene metabolism of NPL1 plasmid from Pseudomonas putida strain BS202."			
RT	Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.			
RM	[6]			

RP SEQUENCE FROM N.A.  
 RC SPECIES=P.putida; STRAIN=ATCC 17484;  
 RA Hamann C.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
 RC SPECIES=P.putida; STRAIN=NCIB 9816;  
 RX MEDLINE=98298434; Pubmed=9634695;  
 RA Kauppi B., Lee K., Carredano E., Parales R.E., Gibson D.T., Eklund H.,  
 Ramaswamy S.;  
 RT "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene  
 1,2-dioxygenase.";  
 RL Structure 6:571-586 (1998).  
 CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
 MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
 BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
 NAPHTHALENE DIHYDRODIOL.  
 CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-  
 dihydronaphthalene-1,2-diol + NAD(+).  
 CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
 CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES  
 CONVERSION OF NAPHTHALENE TO SALICYLATE. AND CATABOLISM OF  
 DIBENZOFIOPHENE (DBI) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO  
 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBI IS LIMITED TO  
 OXIDATION OF THE AROMATIC RING.  
 CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME  
 SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON  
 SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED  
 OF FERREDOXIN REDUCTASE (NDO) AND FERREDOXIN (NDOA) AND ISP IS  
 COMPOSED OF AN HEXAMER OF THREE LARGE ALPHA SUBUNITS (NDOB) AND  
 THREE SMALL BETA SUBUNITS (NDOC).  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 DIOXYGENASE ALPHA SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M29314; AAB47591.1; -;  
 DR EMBL; U49496; AAA92141.1; -;  
 DR EMBL; M60405; AAA16125.1; -;  
 DR EMBL; M83949; AAA25902.1; -;  
 DR EMBL; AF010471; AAB62707.1; -;  
 DR EMBL; AF004284; AAB61373.1; -;  
 DR PIR; JS0071; JS0071.  
 DR PIR; B49343; B49343.  
 DR PDB; 1NDO; 23-MAR-99.  
 DR InterPro; IPR001663; Ring-hydroxyl\_A.  
 DR pfam; PF00848; RING-HYDROXYL\_A; 1.  
 DR PRINTS; PR00090; RING-HYDROXYL\_ALPHA; 1.  
 DR PROSITE; PS00570; RING-HYDROXYL\_ALPHA; 1.  
 KM Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
 KT Dioxygenase; NAD; Plasmid; 3D-structure.  
 FT METAL; 81; IRON-SULFUR (2FE-2S).  
 FT METAL; 83; IRON-SULFUR (2FE-2S).  
 FT METAL; 101; IRON-SULFUR (2FE-2S).  
 FT METAL; 104; IRON-SULFUR (2FE-2S).  
 FT METAL; 208; IRON.  
 FT METAL; 213; IRON.  
 FT METAL; 362; IRON.  
 FT VARIANT; 4; N -> K (IN STRAIN G7).  
 FT VARIANT; 12; S -> F (IN STRAIN ATCC 17484).  
 FT VARIANT; 15; S -> T (IN STRAIN G7).  
 FT VARIANT; 32; K -> R (IN STRAIN G7).  
 FT VARIANT; 50; A -> S (IN STRAIN G7).  
 FT VARIANT; 70; N -> S (IN STRAIN G7).  
 FT VARIANT; 90; SV -> NA (IN STRAIN G7).  
 FT VARIANT; 91;

FT VARIANT 122 122 D -> E (IN STRAIN G7).  
 FT VARIANT 173 173 M -> I (IN STRAIN G7).  
 FT VARIANT 225 225 S -> A (IN STRAIN G7).  
 FT VARIANT 225 225 S -> C (IN STRAIN BS202).  
 FT VARIANT 232 232 A -> V (IN STRAIN G7).  
 FT VARIANT 275 275 A -> S (IN STRAIN G7).  
 FT VARIANT 391 391 E -> R (IN STRAIN G7).  
 FT VARIANT 421 421 Q -> R (IN STRAIN ATCC 17484).  
 FT VARIANT 434 434 H -> D (IN STRAIN G7).  
 SO SEQUENCE 449 AA; 49607 MW; 1FD2F4229664F7A8 CRC64;  
 Query Match 99.8%; Score 2404; DB 1; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 1.5e-181;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MNYNKTLIVSSGSLQKHLHGDDELFOHLEKTFANWILFTHDLSLTPAGDYVAKKG 60  
 DB 1 MNYNKTLIVSSGSLQKHLHGDDELFOHLEKTFANWILFTHDLSLTPAGDYVAKKG 60  
 QY 61 IDEVYSRONGSIRAFLLNCRHGRKTLVSVEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
 DB 61 IDEVYSRONGSIRAFLLNCRHGRKTLVSVEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
 QY 121 KDLYGESLKKCCLCLEKVARVESFHGFIYGCPOEAPPLMDYLGDAWYLEPFRHSGGL 180  
 DB 121 KDLYGESLKKCCLCLEKVARVESFHGFIYGCPOEAPPLMDYLGDAWYLEPFRHSGGL 180  
 QY 181 ELVBPBGVYVKKAMKPAENFVGDAYHGVMTASSLRSGESFFSSLAGAALPPRGAGL 240  
 DB 181 ELVBPBGVYVKKAMKPAENFVGDAYHGVMTASSLRSGESFFSSLAGAALPPRGAGL 240  
 QY 241 QMTSKYSGMGVLMDGYSGVSHADVLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
 DB 241 QMTSKYSGMGVLMDGYSGVSHADVLPELMAFGAKOERLNKEIGVRRARIYRSHLNCYV 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVTVTAIVKMPEDIKRRLADSVORTLGPAGFWS 360  
 DB 301 FPNNSMLTCSGVFKVWNPIDANTTEVTVTAIVKMPEDIKRRLADSVORTLGPAGFWS 360  
 QY 361 DDNNMFTASONGKRYKOSRSDLSNLGFGEDYGYGAVPVGVKSAIGETSRGFYRAY 420  
 DB 361 DDNNMFTASONGKRYKOSRSDLSNLGFGEDYGYGAVPVGVKSAIGETSRGFYRAY 420  
 QY 421 QAHVSSNMAEFERHASTWTELTCTDR 449  
 DB 421 QAHVSSNMAEFERHASTWTELTCTDR 449  
 RESULT 2  
 ID NDOB\_PSEFL STANDARD; PRT; 449 AA.  
 AC 007824;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Naphthalene 1,2-dioxygenase alpha subunit (Ec 1.14.12.12) (Naphthalene  
 1,2-dioxygenase ISP alpha).  
 GN NDOB OR NDOC2.  
 OS Pseudomonas fluorescens.  
 OC Plasmid.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 17483;  
 RA Hamann C.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
 MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
 BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
 NAPHTHALENE DIHYDRODIOL.  
 CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-



```

CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF004283; AAB61370.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske-.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske_1.
CC Pfam: PF00848; Ring_hydroxyl_A_1.
CC PRINTS: PR00090; RINGDIOXGNASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA: 1.
CC KM Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxygenase; NAD; Plasmid.
FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 208 208 IRON (BY SIMILARITY).
FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 362 362 IRON (BY SIMILARITY).
SQ SEQUENCE 449 AA; 49512 MW; 186A7924AD437A1C CRC64;

Query Match 96.3%; Score 2318; DB 1; Length 449;
Best Local Similarity 95.5%; Pred. No. 8; 8e-175;
Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

```

```

DB 361 DDNDNMETASONGKKYQSDLSLNLGFGKDYGDVAVPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSSNAAFEEDASSTWHTLTKTTDR 449
DB 421 QAHVSSSNAAFEEDASSTWHTLTKTTDR 449

RESULT 3
NDOB_PSEAE STANDARD: PRT: 449 AA.
ID NDOB_PSEAE
AC 05194;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DE 15-JUN-2002 (rel. 41, Last annotation update)
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene
DE 1,2-dioxygenase ISP alpha).
GN NDOB OR PAHA3.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAK1;
RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,
RA Kiyohara H.;
RT "The molecular analysis of an NAH7-type gene cluster, pah, located on
RT the chromosome of Pseudomonas aeruginosa PAK1."
RL Submitted (Apr-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)
CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF
CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-
CC NAPHTHALENE DIHYDRODIOL.
CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) = (1R,2S)-1,2-
CC dihydronaphthalene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES
CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF
CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO
CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO
CC OXIDATION OF THE AROMATIC RING.
CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME
CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON
CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED
CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS
CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT
CC (NDOC).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D84146; BAA12240.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske-.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske_1.
CC Pfam: PF00848; Ring_hydroxyl_A_1.
CC PRINTS: PR00090; RINGDIOXGNASE.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA: 1.
CC KM Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxygenase; NAD.
FT METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 208 208 IRON (BY SIMILARITY).

```

FT	MENTAL	213	213	IRON (BY SIMILARITY).
FT	MENTAL	362	362	IRON (BY SIMILARITY).
SEQ	SEQUENCE	449 AA:	49715 MW:	35a1891367222a21C CR664:
Query Match 92.2%; Score 2219; DB 1; Length 449;				
Best Local Similarity 89.8%; Pzed. No. 5.3e-167;				
Matches 403; Conservative 23; Mismatches 23; Indels 0; Gaps 0;				
QY	1	MNYNNKIIIVSESGLSOKHLIHGDEELFOHEHLKTIIFARNMLFLTHDSLIPAGGVYAKMG	60	
DB	1	MYTKNNLVSESGLTQKHLIHGDEELFOHELEITIFARNMLFLTHDSLIPSGDYATKMG	60	
QY	61	IDEVIVSRONDSIRAFILVNCRRHCKTLTVSVENAKGVCVSYHMGFCFSGNGLQSVPE	120	
DB	61	VDEVIVSRQNDGSIIRAFILVNCRRHCKTLTVHAENAKMGVCVSYHMGFCFANGELQSVPE	120	
QY	121	KULYGEGLKCKLGLKEVARVSEFHCFTYCGEDQEAPEPLMDYLGDAAWLTLPYFKHSGCL	180	
DB	121	KELYGALDKCKMGKLEVARVSEFHQFICGDEEAPSLKDYGDGAWLTLPYFKHSGCL	180	
QY	181	ELVGPFGKVIIRANKKAPAPENFVGDAYHYHGWMTFASLSGSEIFSSLAGNALPPEGAGL	240	
DB	181	ELIGPFGKVIIRANKKAPAPENFTGDYHYHGWMTFASLSRGSQVFSFSLAGNALPPEGAGL	240	
QY	241	QMTSKYSGMGVLMWDYSGVHSADLVPELMAFGAKQOEELNKEIGDVARIRIYRSLNCTV	300	
DB	241	QMTSKYSGMGVLMWDYSGVHSADLVPELMAFGAKQOEELNKEIGEVARIYRSHLNCIV	300	
QY	301	PNNNSMLTCSGFYKWNPLIDANTTEWYTAIYEKMDPELKRRLADYVORTLGGPAGFWS	360	
DB	301	PNNNSFLTCSGFYKWNPLIDANTTEWYTAIYEKMDPELKRRLADYVORTGPGAFWS	360	
QY	361	DDNDNNETASQNGKRYKQSHSDLSNLGFGEDVYGDVYGVYKSAIETSYRGFYAY	420	
DB	361	DDNDNNETYSQNAKKYKQSHSDLSNLGFGGVYGDVYGVYKSAIETSYRGFYAY	420	
QY	421	GAHVSSNNAEFEFHASTWHTELTKTTDR	449	
DB	421	GAHVSSNNAEFEEDVSKNWHTELAKTTDR	449	
RESULT 4				
HCAE_ECOLI STANDARD: PRT: 453 AA.				
AC	Q47139: P77590; P78203;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	3-phenylpropanoate dioxygenase alpha subunit (EC 1.14.1.-)			
DE	(Dioxygenin alpha subunit).			
GN	HCAE OR PHDC1 OR HCAE OR HCAAI OR DIGA OR B2538 OR Z3809 OR ECS3404.			
OS	Escherichia coli, and			
OS	Bacteriota; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12;			
RA	Turlin E., Gasser F., Biville F.;			
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=K12 / MG1655;			
RX	MEDLINE=97426617; PubMed=9278503;			
RA	Blaetter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,			
RA	Riley M.J., Collins-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,			
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,			
RA	Mau B., Shao Y.;			
RT	"The complete genome sequence of Escherichia coli K-12.";			
RL	Science 277:1453-1474(1997).			
LN	[3]			
RP	SEQUENCE FROM N.A.			

RC STRAIN=12;  
RX MEDLINE=97349980; PubMed=9205837;  
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,  
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsushashi N.,  
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,  
RA Ohshima T., Oyama S., Saito N., Sampel G., Satoh Y., Sivasubram S.,  
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
RA Yamagata S., Horiiuchi T.;  
RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*  
RT K12 genome corresponding to 50.0-68.8 min on the linkage map and  
RT analysis of its sequence features.";  
RT DNA Res. 4:91-113(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=210744935; PubMed=11206551;  
RA Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfist G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobleck E.J., Davis N.W., Lim A., Dimantanta E.T., Potamouls K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Balthartman T.S.;  
RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
RT Nature 409:529-533(2001).  
RL Nature 409:529-533(2001).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;  
RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
RT DNA Res. 8:11-22(2001).  
RL DNA Res. 8:11-22(2001).  
CC -1- FUNCTION: PART OF MULTICOMPONENT 3-PHENYLPROPYONATE DIOXGENASE.  
CC CONVERTS 3-PHENYLPROPYONIC ACID (PP) INTO CIS-3-(3-CARBOXYETHYL)-  
CC 3,5-CYCLOHEXADIENE-1,2-DIOL (PP-DIHYDRODIOL).  
CC -1- COFACTOR: PROBABLY BINDS A ZFE-25 GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: 3-phenylpropionic acid catabolism.  
CC -1- SUBUNIT: THIS DIOXGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE  
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (HOAE AND HOAF), A  
CC FERREDOXIN (HOAC) AND A FERREDOXIN REDUCTASE (HCAO).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
CC DIOXGENASE ALPHA SUBUNIT FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z37966; CA86018.1; -;  
DR EMBL: AE000340; AAC75591.1; -;  
DR EMBL: D90883; BAA16433.1; -;  
DR EMBL: D90884; BAA16441.1; -;  
DR EMBL: AE005484; AAG57651.1; -;  
DR EMBL: AP002562; BAB36827.1; -;  
DR HSSP: P23094; INDO.  
DR Ecogen: EG13456; hcae.  
DR InterPro: IPR001281; Rleske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rleske; 1;  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RNSDIOXGNSE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
KM Dioxigenase; NAD; Complete proteome.  
FT METAL 85 85 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 87 87 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 105 105 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 105 105 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

FT	METAL	108	108	IRON-SUPFUR (2FE-2S) (BY SIMILARITY).
FT	METAL	213	213	IRON (BY SIMILARITY).
FT	METAL	218	218	IRON (BY SIMILARITY).
FT	CONFLICT	20	20	V -> A (IN REF. 1).
FT	CONFLICT	384	453	GHRAANSKLCLEMGGKRRDDGIGITNTYFETAAAGM YORMDLLSSEMOVLKDTAAYOEVAK -> AATAPTAN CYWKKGLYRKSATATTAFLALLIISFOKLPLEVCTNAGPIE (IN REF. 1).
FT	SEQUENCE	453 AA:	51109 MM;	02535BF5F47643FD CRC64;
SQ	Query Match	30.6%;	Score 737;	DB 1; Length 453;
	Best Local Similarity	36.4%;	Pred. No. 1.7E-51;	
	Matches 165;	Conservative 64;	Mismatches 174;	Indels 50; Gaps
QY	20	IHGDEELFOHELKTIFARNMLFLTHDSLIPABGDYATKMGIDENVLVSRODGSIRAPLN	79	
DB	24	IYTDPIYOLIERIFERIGKMLFLHRSQIPKPGDFNTWGDAYVYVQKQGSIKAFLN	83	
QY	80	VORHNGKTLVSEADNNAKGVCSTYMGWFGSGELOSVPFEKDLIGESI.NKKCLGIKEVA	139	
DB	84	QCRHNAHRVSYADCGNTRAFETCPYHGWSYIGINGELIDVLEBRAYPOGLCKSHWGLINEVP	143	
QY	140	RSESEHGFIYCFDDEAPPLIMDYLDAAAYLEPMF-KHSGGJELIYVPGKVVYIKANWKAP	198	
DB	144	CYESIKGLIFGWMDSAGELRDYLDGMLDRRGSGIETIYGGVQKVIYNCKWKEP	203	
QY	199	AENFYGDAYHVGWTHASSL-----RSGESIFSLGNALP-----	234	
DB	204	AEQFASDQYHALFSHASANQYLGAKDDGSDKRLGD-----GQIARPVWETAKDALQNG	256	
QY	235	--PEBAGLOMPSKYSGSGMYLMDGYSGVSHADLVLELMAFGAKQERLKEIGDVARLY	292	
DB	257	QDGHSGEFTFEKRPANVWV--DGAVSSTYRETYAE-----AEQRIGEYVALRL	303	
QY	293	RSHLCTCFPNNSMLTCSGVFKMNPDIANTYEWVTYATVEKDMPEDKRLRLADSVQRTL	352	
DB	304	AGHNN--IFFTSLMNLGATLRLVWHPRGPDQYEVAAFCITTDAAASDEYKAALENSATRAE	361	
QY	353	GPAGWESDDNDNMETASQNCRRKYOSRSDLSLNLGFGEDVGDVAVPYPGVKSAGIETS	412	
DB	362	GPAGLEQDDSENNWCEIQKLKGHRAHNSKLCLEMGLEGKRRDGIPIIT-NYIFSETA	420	
QY	413	YRGFYRAYQAHVSSSSNAAFEFHASSTWHTELTK	445	
DB	421	AKGMYQRMADLLSSEMOEVLDKTAAAYOEVAK	453	
RESULT 5				
BDL1_PSEPU	BDL1_PSEPU	STANDARD;	PRT;	450 AA.
AC	Q07944;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3).			
GN	BEDC1.			
OS	Pseudomonas putida.			
OC	Plasmid pHM712.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OX	NCBI_Taxid=303;			
RN	(1)			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.			
RC	STRAIN=ML2;			
RX	MEDLINE=93345820; PubMed=8344526;			
RA	Tan H.-M., Tang H.-Y., Joannou C., Abdel-Wahab N.H., Mason J.R.;			
RT	"The Pseudomonas putida ML2 plasmid-encoded genes for benzene			
RL	dioxygenase are unusal in codon usage and low in G+C content";			
CC	gene 130:33-39(1993).			
CC	-1- CATALYTIC ACTIVITY: Benzene + NADH + O(2) -> cis-1,2-			
CC	dihydrobenzene-1,2-diol + NAD(+).			
CC	-1- COFACTOR: PROBABLY BINDS A ZFE-2S GROUP AND AN IRON ATOM.			
CC	-1- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.			

CC	-1	SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BEC1 AND BEC2), A FERREDOXIN (BED) AND A FERREDOXIN REDUCTASE (BEDA).
CC	-1	SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on ways to use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC		-----
DR	EMBL	AF18496; AAL1758.1; .
DR	HSSB	P23094; INDO.
DR	InterPro	IPR001281; Rieske.
DR	InterPro	IPR001663; Ring_hydroxyl_A.
DR	Pfam	PF00355; Rieske; 1.
DR	Pfam	PF00848; Ring_hydroxyl_A; 1.
DR	PRINTS	PR00090; RINGDIOXGNASE.
DR	PROSITE	PS00570; RING_HYDROXYL_ALPHA; 1.
SW		Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron: Dioxygenase; NAD; Plasmid.
KW	DIOXYGENASE	
FT	METAL	96
FT	METAL	96
FT	IRON-SULFUR (2FE-2S)	(BY SIMILARITY).
FT	IRON-SULFUR (2FE-2S)	(BY SIMILARITY).
FT	METAL	116
FT	METAL	116
FT	IRON-SULFUR (2FE-2S)	(BY SIMILARITY).
FT	METAL	222
FT	METAL	222
FT	IRON (BY SIMILARITY)	.
FT	IRON (BY SIMILARITY)	.
FT	IRON (BY SIMILARITY)	.
SO	SEQUENCE	450 AA; 51108 MW; 1BCD5E6AC4CF22C8 CRC64;

Query Match	27.1%	Score 652.5	DB 1:	Length 450;
Best Local Similarity	34.6%	Pred. No. 7.4e-44;		
Matches 148;	Conservative 69;	Mismatches 174;	Indels 37;	Gaps 11;
QY	20	IGHDEELFOHELKIIFARNMLFLTHDSLIPAGDYVYAKMGIDEIVYSRONDSIRAFIN	79	
Db	35	IYTDIEDYQLELEVERFANSMWLLGHETHIRKRGDYFTTYMGEDPVVVYRKQKASIAVLEFN	94	
QY	80	VCRHNGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEFKDLYGESLNKKGLGKE--	137	
Db	95	QCRHGMKRICRSDAGNAKAFKFCSYHGMAVYDPAAGNLINPYEAESFA-----CLDKKEMS	148	
QY	138	--VARVESFHGIGYCFQCEAPLMDYDGDAMAYEPNF--KHSGLLELVGPPKVIKAN	194	
Db	149	PLKARVETIKGLIFANMDENAIDLTDTYGEAKFYDHMDIKREATEVEIPGIQKVCICN	208	
QY	195	WKPAEENFVGDAVYHGMW--HASSLRSGESIFSSLGNALPPEGAGLOMTSKYSGSGMGLV	253	
Db	209	WKPAAEGQCSMDYHAGCTIAHLISGIIAGLPEDLELDLA--PP-----KFGQYRAS	257	
QY	254	WDGY--SGVHSAD-----LYPELMAF--GGAQOERLNKEIGDVR--ARIYRSHLNTVF	301	
Db	258	WGGHSGSGEYIDDPNMMLAMGPKVYSYLTGEPAAEKAERGSIERGTRKIMLEHM--TVF	315	
QY	302	PNNSLMTCSSGFYKWNPIIDANTTEWYTAIYEKMDPELCKRLRLDSVQRTLGPAQFWESD	361	
Db	316	PTCSFLPELVNTIRTWHPGPNVEVEWATVVDADAPDDIKKEEPRQRLTRTSAGGVFROD	375	
QY	362	DNDMMETASONGKRYKQSRDLSLSNLGEEGVYGGDAVYPGVGKSAIGETSYRGFYRAYQ	421	
Db	376	DGEWMLVQLHLRHKAKASRPFNEMSMGQYVDNDPIYGRISNNVYSEELARGIYAHML	435	
QY	422	AHYSSSWY	429	
Db	436	KMMTSPDW	443	

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toluene 2,3-dioxygenase alpha subunit (EC 1.14.12.-).
GN
OC Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC STRAIN=FL;
RX MEDLINE=89359301; PubMed=2670929;
RA Zylstra G.J., Gibson D.T.;
RT "Toluene degradation by Pseudomonas putida FL. Nucleotide sequence of
  the todC1C2Bade genes and their expression in Escherichia coli.";
RL J. Biol. Chem. 264:14940-14946(1989).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: Toluene degradation; first step.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
  TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (TODC1 AND TODC2), A
  FERREDOXIN (TODB) AND A FERREDOXIN REDUCTASE (TODA)
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
  DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
CC EMBL: J04996; AAA26005.1; -.
CC PIR: A36516; A36516.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00335; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 IRON (BY SIMILARITY).
FT METAL 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 50944 MW; 038C80F197F3485D CRC64;
Query Match 27.0%; Score 650.5; DB 1; Length 450;
Best Local Similarity 34.4%; Pred. No. 1,1e-43;
Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;
QY 20 IHDEDELQHELTIFARNMLFTHDSLIPADQVYAKMGIDEIYVSRODGSIRAFLN 79
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 35 IYDEDEDLYOLELVEFARSMLLGHEFOIKRPQDVIITYTMEDDPVYVYKQDASIVFLN 94
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 80 VCRHGRKTLVSVEGNAKNGVCSYHGWGFGSGNGELQSVPEKDLYGESLNKCKGLKE- 137
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 95 QCRHGRICRADGNAKAFCSYHGMAVDPAAGMLVNVVPEAESFA-----CLNKREMS 148
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 138 --VAARVESFGITGCDQEPRLMDYLGAAMWLEPMF-KHSGLELVPPGKVVYKAN 194
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 149 PLKARVETTYGLIFAMWDENAVLDITYLEKAKFYMDHMLDTEGTEAIPVQKWVIFCN 208
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 195 WKPAEVEFVDAYHVGMT-HASSLSRG--ESIFSSLAGNALPPEGAGLQMTSYSGSGMG 251
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 209 WKRAAEQFCSDMTHAGTTSILSGILAGLPEDL-----EMADLAPPTYGKQIRASNGHGS 263
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 252 VLWDGYSGVASHADLVPELMAF--GGAKOERLNKEIGDVR-ARIYRSHLNTCTVPPNNSML 307

```

```

DB 264 GFYGDPLMLAIAGPKVTSYWGEPASKEAERLGSVERGSKLMVEHM--TVFPCSFLL 321
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 308 TCSGVFKYMNPIDANTEVMTYALVEKMDPEDLKRLRLADSVQRTLGPAGWESDDNDNE 367
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 322 PGINTVRTMHPRGNEVEWMAFTVVDADAPDDIKKEFRROTLPFRFSAGVFEDDDEENVV 381
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 368 TASQNGKKYQSRDSDLNSLGFGEVDYGDVAPGVCSAIGETSYKGFYRAVQAHVSSS 427
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 382 EIGHLRGKHKRSPFAEKSMQDVTNDPVIYRISNNYSEEARGLYAHMLRMATSP 441
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 428 NW 429
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 442 DW 443
  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
RESULT 7
BPHA_BURCE STANDARD; PRT; 458 AA.
ID BPHA_BURCE
AC P37333;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
  dioxygenase).
GN BPHA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LB400;
RX MEDLINE=92234948; PubMed=1569021;
RA Erickson B.D., Mondello F.J.;
RT "Nucleotide sequencing and transcriptional mapping of the genes
  encoding biphenyl dioxygenase, a multicomponent
  polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain
  LB400.";
RL J. Bacteriol. 174:2903-2912(1992).
RN [2]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RC STRAIN=LB400;
RX MEDLINE=96011369; PubMed=7592331;
RA Haddock J.D., Gibson D.T.;
RT "Purification and characterization of the oxygenase component of
  biphenyl 2,3-dioxygenase from Pseudomonas sp. strain LB400.";
RL J. Bacteriol. 177:5834-5839(1995).
RN [3]
RP ERRATUM.
RA Haddock J.D., Gibson D.T.;
RT "Bacteriol. 178:2158-2158(1996).
  -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
  phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
  -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit.
  -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
  first step.
  -1- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
  Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase
  (Bphg) must be present to obtain activity.
  -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
  DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
CC EMBL: M86348; AAB63425.1; -.
CC PIR: B41858; B41858.
DR

```

```

DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxigenase; NAD.
FT INIT_MER 0
FT METAL 99 99 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 122 122 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 232 232 IRON (BY SIMILARITY).
FT METAL 238 238 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51382 MW; AFSBFLD05BAFDF4 CRC64;

Query Match 25.6%; Score 617.5; DB 1; Length 458;
Best Local Similarity 32.9%; Pred. No. 4,3e-41;
Matches 146; Conservative 69; Mismatches 186; Indels 43; Gaps 12;

QY 11 ESLSKHLIHDELFQHELTIFARWMLFTHDSLIAPAGDYATKAGIDEVYTSRON 70
DB 30 EKGLDPR-IYADQSLYLELRFVGRSWMLLGHESHVETGDFLATYMGEDPVVVRQK 88
QY 71 DGSIRAFVWCRHGRKTLVSVEAGNAKGFVSGHGFSGNGLSVPEKDLGSESLNK 130
DB 89 DSIKFTLMOCHRRGKRIRGRSDGNAKAFCTSYHGAYIAGLKVVPPEKEKFCCKKG 148
QY 131 KCLGKEV-----ARVESFHGFIYGCDFQDEAPRLMDYLGDAAWYLEPMEKHS-GGLELV 183
DB 149 DC-GFDKAEWGLPQARVATYKGLVFMMDVQAPDLETYIGDARPYDMLDRPAGTVAI 207
QY 184 GPPGKYVITANKAPAEENVGDAIHVG-WTHASSLSRSGSIFSSLAGNAALPPEG--AGL 240
DB 208 GGNQKVVIVICNMKFAEEOCSDMYHAGTTHLSGLIAG-----IPEMDLSQA 255
QY 241 QMTSKYSGMGVLMDGYSVSHADLVPELMARFSAK-----OERLNKEID 286
DB 256 QLPFK-GNCFRAWGGHSGWYVDDEPGSLIAYWGPVYOTYTWEGRAELAEORLGHITGAP 314
QY 287 VRARIYRSHLNCVFPNNMLTCSGVFKVWNPIDANTTEVYATIAVEKMDPEDLKRLAD 346
DB 315 VR-RWVGQHM--TIFPTCSFLPTFNIRIMHPRGRVEIEMVAFITLDADAPAEIKETERR 371
QY 347 SVQRTIGPAGFESDNDNMETASONGKRYQSRDLSNLGFGEDVYGDAYIPGVYGS 406
DB 372 HNIIRNSAGGVFEODDGEWVEIQKLRGYKAKSOPLNOMGLGRSQTGHPRDFPGWVG-Y 430
QY 407 AIGETSYRGFYRAYOAHVSSNMA 430
DB 431 VVAEEARGMHYHMMRMSEPSWA 454

RESULT 8
BPHA.COMTE
ID BPHA.COMTE STANDARD: PRT: 457 AA.
AC 046372;
DB 15-JUL-1998 (Rel. 36, Created)
DB 15-JUL-1998 (Rel. 36, Last sequence update)
DB 15-JUN-2002 (Rel. 41, Last annotation update)
DB Biphényl dioxigenase alpha subunit (Ec 1.14.12.18) (Biphényl 2,3-
DE dioxigenase).
GN BPHA.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; Beta subdivision; Comamonadaceae; Comamonas.
NCBI_TaxID:285;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-B-356;
RA MEDLINE-97043812; PubMed-8890734;
RA Sylvestre M., Stiois M., Hurtubise Y., Bergeron J., Ahmad D.,

```

```

RA Shareck F., Barriault D., Guillemette I., Juteau J.M.;
RT "Sequencing of Comamonas testosteroni strain B-356-
RT biphenyl/chlorobiphenyl dioxigenase genes: evolutionary relationships
RT among Gram-negative bacterial biphenyl dioxigenases."
RL Gene 174:195-202(1996).
CC -I- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) -> (2R,3S)-3-
CC phenyleyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -I- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -I- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -I- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphn subunits. A ferredoxin (Bphf) and a ferredoxin reductase
CC (Bphg) must be present to obtain activity (By similarity).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGNASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U47637; AAC44526.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rieske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rieske; 1.
DR Pfam; PF00848; Ring_hydroxyl_A; 1.
DR PRINTS; PR00090; RINGDIOXGNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxigenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 457 AA; 51691 MW; DI33FC0635FACBF5 CRC64;

Query Match 25.6%; Score 617; DB 1; Length 457;
Best Local Similarity 33.1%; Pred. No. 4,7e-41;
Matches 144; Conservative 73; Mismatches 172; Indels 46; Gaps 13;

QY 20 INGEDELFGHELTIFARWMLFTHDSLIAPAGDYATKAGIDEVYTSRONDSIRAPLN 79
DB 39 IYADQDLYOLELRFVGRSWMLMGLHETHIPIKIGDYLTYYMGEDPVIMVRQKDSIKVPLN 98
QY 80 VCHNRKTLVSVEAGNAKGFVSGHGFSGNGLSVPEKDLGSESLNKKCLGKEV- 138
DB 99 QCRHGRMIRVSDGNKAKAFCTYHGMAVDIAGNLVNVPEKFAFCDDKRGCD-GFDKAD 157
QY 139 -----ARVESFHGFIYGCDFQDEAPRLMDYLGDAAWYLEPMEF-KHSGLELVGPPGVYIK 192
DB 158 WGPLQARVETKYGLVAMNDPEARPDLYTSLDAMPYMDVMDLRTETAGTAIGTIQKWVIP 217
QY 193 ANMKAPAEENVGDAIHVG-WTHASSLSRSGSIFSSLAGNAALPPEGAGLOMT---SKY 247
DB 218 CNMKFAEEOCSDMYHAGTTHLSGLIAG-----LPPE---MDLTQIQLSKNG 262
QY 248 SGNGVLMIDGY-----SGVSHADLVPELMAR--GGAGQERLNKKEIGYR-ARIYRSH 296
DB 263 NQFRSAWGGHAGWFTNDSILSLSVGPKITQYTWIGQPAEAKKARVPOPLIIMDGQHM 322
QY 297 NCTVFPNNMLTCSGVFKVWNPIDANTTEVYATIAVEKMDPEDLKRLADSVORTLGAPG 356
DB 323 --TVFPTCSFLPGINTIRWHPGRVNEVWMAFLVDADAPEDIKKEFPLQNIIRTFNAG 380
QY 357 FWSDDNDNMETASONGKRYQSRDLSNLGFGEDVYGDAYIPGVYGSAA--IGETSYR 414

```

```

Db 381 VFEDDDEENWEIQRVMGHRKAKSTSLCAKMGILNPNKNNDPYP---GKTAVVAEENAR 437
QY 415 GFYRAYQAHVSSNMK 429
| : : : :
Db 438 GMYHHMSRMHSEPSM 452

RESULT 9
BPHA_PSEPS
ID BPHA_PSEPS STANDARD; PRT; 458 AA.
AC 052028;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (biphenyl 2,3-
DE dioxygenase).
GN BPHA OR BPHAL.
OS Pseudomonas pseudoalcaligenes.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=330;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KF707;
RX MEDLINE=92165849; Pubmed=1537863;
RA Taira K., Hirose J., Hayashida S., Furukawa K.;
RT "Analysis of bph operon from the polychlorinated biphenyl-degrading
RT strain of Pseudomonas pseudoalcaligenes KF707.";
RL J. Biol. Chem. 267:4844-4853(1992).
CC -1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (Bphf) and a ferredoxin reductase
CC (Bphn) must be present to obtain activity (By similarity).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M83673; AAA25743.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rleske.
DR InterPro: IPR001663; Ring-hydroxyl_A.
DR Pfam: PF00355; Rleske.1.
DR Pfam: PF00848; Ring-hydroxyl_A.1.
DR PRINTS: PR00090; RINGDIOXYGENASE.
DR PROSITE: PS00570; RING-HYDROXYL_ALPHA.1.
KW Aromatic hydrocarbon catabolism, Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 100 100 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 233 IRON (BY SIMILARITY).
FT METAL 239 239 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51437 MW; A6123AD09F97E462 CRC64;

Query Match 25.68; Score 617; DB 1; Length 458;
Best Local Similarity 33.18; Pred. No. 4,7e-41;
Matches 147; Conservative 70; Mismatches 183; Indels 44; Gaps 13;

QY 11 ESGLSQKHLHGDELFQHEHLEKTFARNWLFJLTHDSLIPAGDGYVYAKMGIDEIVSRQN 70
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 31 EKGILDPK-IYADOSLYELEERVFGSRWLLGHESHVPETGFLATYMGEDPVVWVRQK 89
QY 71 DGSIRAFNLVCRHREKTLVSVEAGNAGKGFVCSYGMWFGSGNGELQSPFEKLDYGESLTK 130
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 90 DKSTIVFLNOCRRHGRMRCSRSDAGNAKAFCTSYGMATVDIAGKLVNPFKEAFCDKREG 149
QY 131 KCLGLEVEY-----ARVESFHGFTYGCDFDEAPPLMDYLDPAWYLEPRKHS-GGLELY 183
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 150 DC-GFDKAEWGPRLQARVATYKGLVFANMDVQAPDLLEYLDGARYMDVMDLRIPAGTVAI 208
QY 184 GPPKGVITKANWKAFAENFVGDAYHVG-WTHASSLRSGESIFSSLAGNALPPDG--AGL 240
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 GGMOKWVTPCWKKPAAEQFCSDMYHAGTMSHLSGILG-----MPPEMLSHA 256
QY 241 QMTSKYSGMGVLWDGYSGVHSADLVPELMAFGAK-----QELNKEIGD 286
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 QVPTK-GNQFRAGWGHSQGFVDEPGMLAAVMGPKVTQYWTGSPAADLAEOIRGHTM-P 314
QY 287 VRARIYRSHLCTVFPNNSMITCSGVFKVWNPIDANTTEWYAIYEKMPEDLKRRLAD 346
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 315 VR-RMFQGMH--SVFPICSLPALNTIRTHRRGPNLEIWMATFLVDADAPAIKEEYR 371
QY 347 SVQRTLPAGFWEESDDNDNMTASONGKRYOSRSDLLSNLGFGEEDYGDVYPGVYGS 406
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 372 HNIRTFSGAGVFEDDDEENWEIQGLRGYKAKSQPLNAQMGIGRSQGTGHPDEPVNG-Y 430
QY 407 AIGETSYRGFRAVQAHVSSNMA 430
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 431 VYAEBAARGMTHWMRMSEPSMA 454

RESULT 10
BNZA_PSEPU
ID BNZA_PSEPU STANDARD; PRT; 448 AA.
AC P08084;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3) (p1 subunit).
GN BNZA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BE-81.
RX MEDLINE=88032840; Pubmed=3667527;
RA Irie S., Doi S., Yorifuji T., Takagi M., Yano K.;
RT "Nucleotide sequencing and characterization of the genes encoding
RT benzene oxidation enzymes of Pseudomonas putida.";
RL J. Bacteriol. 169:5174-5179(1987).
CC -1- CATALYTIC ACTIVITY: Benzene + NADH + O(2) = cis-1,2-
CC dihydrobenzene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BNZA AND BNZB), A
CC FERREDOXIN (BNZC) AND A FERREDOXIN REDUCTASE (BNZD).
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17904; AAA25735.1; -.
DR PIR: A29830; A29830.
DR HSSP: P23094; INDO.

```

Query Match	Best Local Similarity	Score	DB 1	Length	448:
Matches 142;	Conservative 67;	Mismatches 186;	Indels 27;	Gaps 11;	
DR	InterPro: IPR001281; Rieseke				
DR	InterPro: IPR001663; Ring_hydroxyl_A				
DR	Pfam: PF00355; Rieseke; 1				
DR	Pfam: PF00848; Ring_hydroxyl_A; 1				
DR	PRINTS: PF00090; RINGDIOXGNASE				
DR	PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1				
KW	Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;				
KW	Dioxygenase; NAD.				
FT	METAL	96	IRON-SULFUR (2FE-2S) (BY SIMILARITY)		
FT	METAL	98	IRON-SULFUR (2FE-2S) (BY SIMILARITY)		
FT	METAL	116	IRON-SULFUR (2FE-2S) (BY SIMILARITY)		
FT	METAL	119	IRON-SULFUR (2FE-2S) (BY SIMILARITY)		
FT	METAL	222	IRON (BY SIMILARITY)		
FT	METAL	228	IRON (BY SIMILARITY)		
SQ	SEQUENCE	448 AA; 50208 MW; 601DDB5B2F040BE7 CRC64;			
Query Match	24.4%;	Score 588.5;	DB 1;	Length 448:	
Best Local Similarity	33.6%;	Pred. No. 7.9e-39;			
Matches 142;	Conservative 67;	Mismatches 186;	Indels 27;	Gaps 11;	
QY	10HGDELEFQHELEKTTFARNWLEFLTHDSLIPADGDYVTAKMGIDEIVYSHQNDGSIRAFLN 79				
DB	35 IYTDDELVQLEDERVAFASWMLLGHETQIRKPGDYITTYMGEDPVVVRQKASIAVFLN 94				
QY	80 VQRHKGKTLVSEAGNANGFVCSHYGNGFGSGNGELQSVFPEKDLYGESLNKKCLGKE-- 137				
DB	95 QCRHGMGRICRADAGNNAKAFCTSYHGMAVDYTGNNVNPYEAESEA-----CLNKKEWS 148				
QY	138 --VARSEPHGRTYCFCFOEAPFLMDYDGDAAWYLEPME-KHSGGELUGPPGKVITAN 194				
DB	149 PLKAVERTYKGLIFANWDENAVDLDTTYGEAKFTYMDHMDRTDAGTEALPGYQKWTIPCN 208				
QY	195 WKPAENFVGADYHVGMT-HASSLRSG--ESIFSSLAGNAALPPEGAGLOMTSKYSGWG 251				
DB	209 WKFAEORCSMDYHAGTTSHLGILAGLPEDL-----ENADLAPPTVYGQYRASNGGHS 263				
QY	252 VLMDEYSGVHSDVLEPELMAF--GGAQKQERLKEITGVR--ARIYRSLNCTVFPNNSML 307				
DB	264 GFYVDDPNMLMALIMPKYTSYWTESGPASEKAERLGSVERGSKLWVEHM--TVFPTCSFL 321				
QY	308 TCSGVFKWNPIDATWTTFWMTVAIEKMPEDLKKRLRLADSVORTLGPGAFWESDDNDNNE 367				
DB	322 PCINTVFRLLASRAARGEWMAFTVVDADDPDDIKEEPRAL-RTSPVA-CSSRTTGRTG 379				
QY	368 TASQNGKKYOSRSDSLNSLGGEDYVDGAPVGVAGSAIETSGTGRYRYQAHVSS 427				
DB	380 SRSSVSCATSRFSRFNMEMSMDQIVDNDPVYPGKISNNVYSEEARGLAHMLPMATSP 439				
QY	428 NW 429				
DB	440 DW 441				
RESULT 11					
BPA1_PSES1	STANDARD;	PRT;	458 AA.		
AC	Q52438;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-dioxygenase).				
GN	BPAH1.				
OS	Pseudomonas sp. (strain KKS102).				
OC	Bacteria; Proteobacteria.				
OX	NCBI_TaxID=307;				
RM	(1)				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94324977; Pubmed=8048958;				
RA	Fukuda M., Yasukouchi Y., Kikuchi Y., Nagata Y., Kimbara K.,				
RT	Horikuchi H., Takagi M., Yano K.,				
RT	*Identification of the bphA and bphB genes of Pseudomonas sp. strains KKS102 involved in degradation of biphenyl and polychlorinated				

```

biphenyls.";
RT Biochem. Biophys. Res. Commun. 202:850-856(1994).
CC -i- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -i- COPACINOR: Binds 1 2re-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -i- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -i- SUBUNIT: Heterohexamer consisting of 3 BphA1 subunits and 3
CC BphA2 subunits. A ferredoxin (BphA3) and a ferredoxin reductase
CC (BphA4) must be present to obtain activity (By similarity).
CC -i- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC
CC
DR HSPB: D17319; BAA04137.1; -.
DR HSPB: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PRO0090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 103 103 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 121 121 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 124 124 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 234 234 IRON (BY SIMILARITY).
FT METAL 240 240 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51876 MW; 128E8C38E2A3CFAA CRC64;
-----
Query Match 24.4%; Score 587; DB 1; Length 458;
Best Local Similarity 33.6%; Pred. No. 1,1e-38;
Matches 147; Conservative 65; Mismatches 174; Indels 52; Gaps 14;

QY 20 IHGDELPHOHELTETPARFWLFLTHDSLIRPAGCDVYTKMGIDETIVSRQNDGSTRAPLN 79
DB 40 IYADDDLEIETELERFPAKSWLLGHEAHIPKGDVLTYYMGDPYIMVRQKQSGIKVEFN 99
QY 80 VCRHNGKTLVSEADANAKAGFVCSYHGMGFGSGELOSVPFEKDLGESSLNKKCLGKEV- 138
DB 100 QCRHHGSMICRSDANAKAFCTTYHGMAYDIAGNLVNNPYKEAKCDKKEGQC -GFDRK 158
QY 139 -----ARVESFHGTLTGCFDDEAPRLMDYLGDAAVYLEPMF -KHSGLGELVGPCKVYIK 192
DB 159 WGPLARARVETYLGLTFANWDADEAPDLKTYLSDAMPYMDMLDRTAGTTVVGGMQKWYLP 218
QY 193 ANMKRAPENFVDADVHVG -WTHASSLRSGESIFSSLAGMAALPRPGAGLOMT -----SKYG 247
DB 219 CMMKRAAOFCSDMTHAGTMHLSGVLS -----SLRPP -----MDLTLYOMSKNG 263
QY 248 SGMGLVIMPGYSGVSHADLVPELMAEGAK -----QELNKEIGDVRR -IYRSIL 296
DB 264 SGRFAAMGSHGSGWITINDAAILMAYMGKITQYVWQSGRAEKAARKLQMLPQTMFGDM 323
QY 297 NCTYEPNNMSLTCGSGVFYVMEIDANTTEV -----WTYAIVERKMDPEDLKRLADSVQRTLG 353
DB 324 - -TVPTCSFLPGIWTIRSMHPRGNEVECGPSMSBMR -----PEDIKEEPRQIRTFEN 378
QY 354 PAGFWESDDNNMFEASQNGKKYQSRDSDLNLGFGEDVYGDAAVYPCGVGKSA -IET 411
DB 379 AGGTFEODDGEHWEIQLGRGHRKAKSAPLQOMGLNVPKNSNPDPK - -GXTAVVAAE 435
412 SYRGFYRAYOAHVSSSNW 429

```

Db 436 AARGMYHHMARMSEPNW 453

# RESULT 12

XYLX\_PSEPU STANDARD; PRT; 454 AA.

ID XYLX\_PSEPU

AC P23099;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Toluene 1,2-dioxygenase alpha subunit (EC 1.14.12.-).

GN XYLX.

OS Pseudomonas putida.

OC Plasmid TOL PMW.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas

OX NCBI\_taxid=303;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92155191; PubMed=1740120;

RA Neidle E.L., Hartnett C., Ornston L.N., Bairoch A., Reikik M.,

Harayama S.,

"Cis-diol dehydrogenases encoded by the TOL PMW0 plasmid xylX gene

and the Acinetobacter calcoaceticus chromosomal bend gene are members

of the short-chain alcohol dehydrogenase superfamily."

Eur. J. Biochem. 204:113-120(1992).

-1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.

-1- PATHWAY: DEGRADATION OF BENZOATE AND TOLUATE TO CATECHOL.

-1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS: THE

TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (XYLX AND XYLZ), AND AN

ELECTRON TRANSFER COMPONENT (XYLZ).

-1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

DIOXYGENASE ALPHA SUBUNIT FAMILY.

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

QY 139 ARVESFHGIFCYCPDOEAPRLMDYLGDAMYLEPFMKHS--GGELVGPCKVYIKANMKA 197

Db 152 ARFASVREGFLGSLREDVAPLEEFLLGESRKVYIDWVVDQSEGLEVLRGSTVYEEENWV 211

QY 198 PAENFVGDAHYG--WTHASS-----LR-SGESITSSLAGNALPPEGAGLOMTSKYGS 248

Db 212 QVEN-GADGYHVSIVHNNYATQOQRKLRDAGDDIRAMTA--SSWGDGGGFF--SFEN 265

QY 249 GAGVLMDGYSVSHADLVPELMARFAGAKOEFLNKETIGDVAR--IYRSHLCTVFPNNSM 307

Db 266 GCHWMAWARGPKRPLF-----AERDLASEFGEARADWIGVSRNLCYPLUYIM 317

QY 308 TCSCG-VFKWNPIDANTTEWTVAYVER-DMPEDLKR 342

Db 318 DQFSQLRTRPLSVDRFEITLYIAPKGEPRRARR 354

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----

-----



Query Match	Best Local Similarity	Matches 128:	Conservative 78:	Mismatched 204:	Indels 52:	Gaps 20:									
16.6%, Score 399, DB 1, Length 461;	27.7%, Pred. No. 6,5e-24;														
5 KNTIV--SESGISQKH-LIHGDEEFLFOHEKLTIFARNMFLTHDSILIPARGDYATAKMGI 61	16 DELVADNETGTFKLRHSVFTDQALFLEMKITFEGNWNVYILNHEQIPNNNDYTTYIOR 75	16 DEVIYSRKNDGSIKRAFLVNCRRHGTIVSVEAGNAKGEVSYHGMFGSGNGLQSVPEEK 121	76 QPILIAARRNNGELNMINACSHRGALDGRHKRGNKTTTYPCCPHGWTFFNSGKLTKYKDPSS 135	122 DL-VESLSNKK-CGLKKEVARYESHGFVGGDQEDAPRLMDLIDDAAMLPRMKHSG- 178	136 DAGVSDCNDQSDHDLKVAARESTKGFLEFGLNPDVPSLOGLFETTKITIMIVQSQDQ 195	179 GLEIVPGPKVYIKANKKAPENFVGDAYVHG--WTNASSD-----RSGESIFSLAG 229	196 GLEIVRGVSTYTBEGNKLTAE-N-GADGVHSAVMHNTAAQTQHNKKEKQAGDTIAMSAG 254	220 NAALPBEAGLQMTSKTG--SGMYLMDGYSVHSADLVPELM---AFGAKQERLNKE 283	255 SWG--KHGGG-----SYGFEGHGMILMTQWPNEDRPNPKAEYTEKFGAMSKWM-- 304	284 IGDVARIYRSHLNTCTVPPNNSMLTCSG-VFKNVMPIDANTCEVNTYALVE-KDMPEDLK 341	305 -----IERSR-NICTLPNYILMQFGSQIRVLRPIISNKTKEVITITTCIAPYGEARPEANA 356	342 RRLADSVORTGTPAGEFWESDDDNKNETASQNGKKYQSRDSDLLSLMGEGEDVYG--DA-- 397	357 RRIR-QYEDDFNASSGMARPPDLEEFRAQAGYAGIELEWMDCR--GSKHWIYGPDDAN 413	398 --VYPGVGKSAIGETSYRGFRAYVQAHVSSNMAEPFHAS 436	414 EIGLKPAISGIKTEDEGLYLAOHQYILKMSKQALAEKEFAS 455

```

RX MEDLINE=97251358; PubMed=9097040
RA Itch T., Alba H., Babba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Yasunodanarum S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horikuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40,1-50.0 mln region on the linkage map.";
RL DNA Res. 3:379-392(1996).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / EDL933 / ATCC 700927;
RC MEDLINE=21074935; PubMed=11206551;
RA Pena N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posral G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocus K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=O157:H7 / RIMD 0509952;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobé T.,
RA Iida T., Takami H., Honda T., Saekawa C., Ogasawara N., Yasunaga T.,
RA Kusaba S., Shibata T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -I- COFACTOR: PROBABLY BINDS A 2PE-2S GROUP AND AN IRON ATOM.
CC -I- SUBUNIT: PROBABLE HETERODIMER OF YEAM AND YEAX
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. See http://www.isb-sdb.ch/announce/
CC or send an email to license@sib-sdb.ch).
CC -----
DR EMBL: AE000274; AAC74872.1; -
DR EMBL: D90823; BAAL15597.1; ALT_INT.
DR EMBL: D90824; BAA15606.1; ALT_INIT.
DR EMBL: AE005403; AAG56791.1; -
DR EMBL: AP002558; BAB35934.1; -
DR EcoGene: EG13509; yeaw.
DR InterPro: IPR001281; Rleske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rleske_1.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA_FALSE_NEG.
KW Hypothetical protein, Oxidoreductase, Iron sulfur, Iron, Dioxygenase;
KW NAU: Complete proteome.
FT METAL 89
FT METAL 89 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 91 91 IRON-SULFUR (2PE-2S) (BY SIMILARITY).
FT METAL 109 109 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 211 211 IRON (BY SIMILARITY).
FT METAL 216 216 IRON (BY SIMILARITY).
SQ SEQUENCE 374 AA; 42561 MW; BB386ACA9585606 CRC64;
Query Match 10.4%; Score 251; DB 1; Length 374;
Best Local Similarity 26.9%; Pred.No.2,le-12;
Matches 99; Conservative 53; Mismatches 126; Indels 90; Gaps 19;

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:55:37 ; Search time 13.1667 seconds  
(without alignments)  
3278.305 Million cell updates/sec

Title: US-09-843-250-35

Sequence: 1 MNVNNKILVSESGLSQKHLI.....AEFEHASTWHTLTQTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2404	99.8	449	2 JS0071	naphthalene dioxyg
2	2404	99.8	462	2 S27632	naphthalene 1,2-di
3	2346	97.4	449	2 JN0644	naphthalene 1,2-di
4	2292	96.2	449	2 C55217	polycyclic aromati
5	2054	85.3	447	2 JC5352	2-nitrotoluene dio
6	957	35.7	459	2 T31134	naphthalene dioxyg
7	774	32.1	450	2 T31256	terminal oxygenase
8	737	30.6	453	2 A65031	biphenyl dioxygena
9	737	30.6	453	2 D91054	biphenyl dioxygena
10	737	30.6	453	2 G85898	biphenyl dioxygena
11	725.5	30.1	455	2 T31258	biphenyl dioxygena
12	652.5	27.1	450	1 JN0812	aromatic oxygenase
13	650.5	27.0	450	1 A36516	benzene 1,2-dioxyg
14	639	26.5	461	2 S51757	benzene 1,2-dioxyg
15	617.5	25.6	459	1 B41858	biphenyl dioxygena
16	617	25.6	457	1 J04993	biphenyl dioxygena
17	617	25.6	458	1 A42409	biphenyl dioxygena
18	605	25.1	431	2 JN0098	carbazole dioxygen
19	588.5	24.4	448	1 A29830	benzene 1,2-dioxyg
20	587	24.4	458	2 JC4267	biphenyl dioxygena
21	452	18.8	469	2 T50934	dioxygenase D1AL,
22	416	17.3	471	2 T31281	benzoate 1,2-dioxy
23	410.5	17.0	454	2 A41659	benzoate 1,2-dioxy
24	409.5	16.4	455	2 E83332	toluene 1,2-dioxyg
25	394	16.4	461	2 S23477	anthracene dioxy
26	389.5	16.2	464	2 G83331	probable ring-hydr
27	339.5	14.1	424	2 E83384	biphenyl dioxygena
28	331	13.7	426	2 T31278	aromatic oxygenase
29	317.5	13.2	391	2 T31251	

30	311.5	12.9	468	2 G97447	hypothetical prote
31	311.5	12.9	468	2 AH2665	ring hydroxylating
32	294	12.2	429	2 F82970	probable ring hydr
33	287	11.9	412	2 H95311	probable aromatic-
34	283	11.8	415	2 AG3320	benzoate 1,2-dioxy
35	271.5	11.3	404	2 H87635	Rieske 2Fe-2S faml
36	266	11.0	420	2 T31285	biphenyl dioxygena
37	254.5	10.6	374	2 AF0304	probable dioxygena
38	251	10.4	374	2 C85791	probable choline m
39	251	10.4	374	2 G90942	probable choline m
40	251	10.4	374	2 B64941	probable choline m
41	233.5	9.7	382	2 G70946	probable dioxygena
42	224	9.3	446	2 T14542	choline monooxygen
43	223	9.3	439	2 T09214	choline monooxygen
44	214.5	8.9	186	2 JC5354	2-nitrotoluene dio
45	186	7.7	426	2 T08550	choline monooxygen

## ALIGNMENTS

```
RESULT 1
JS0071
naphthalene dioxygenase (EC 1.14.12.-) ndob protein - Pseudomonas putida
C:Species: Pseudomonas putida
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Aug-1999
C:Accession: JS0071
R:Kurkela, S.; Lehtsaalaho, H.; Palva, E.T.; Teeri, T.H.
Gene 73, 355-362, 1988
A:Title: Cloning, nucleotide sequence and characterization of genes encoding naphthal
A:Reference number: JS0070; MUID:89211973; PMID:3243438
A:Accession: JS0071
A:Molecule type: DNA
A:Residues: 1-449 <KUP>
A:Cross-References: GB:M2914; NID:9151392; PIDN:AAB47591.1; PID:9151394
C:Comment: Naphthalene dioxygenase system is composed of three proteins.
C:Genetics:
A:Gene: ndob
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein
F:71-119/Domain; Rieske [2Fe-2S] homology <RSK>
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status
Query Match 99.8%; Score 2404; DB 2; Length 449;
Best Local Similarity 99.8%; Pred. No. 1-2e-183;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	1	MNVNNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFLTHDSLIPAPDDYVAKMG	60
DB	1	MNVNNKILVSESGLSQKHLIHGDELFQHELKTIFARNMLFLTHDSLIPAPDDYVAKMG	60
QY	61	IDEVYVRONDGSTRFLNVCRRKGTIVSVAGNAKGVCSYHGFGSGNELSVPE	120
DB	61	IDEVYVRONDGSTRFLNVCRRKGTIVSVAGNAKGVCSYHGFGSGNELSVPE	120
QY	121	KOLYGSLNKKCLGLKEVARVSPFHGFTYCGPDEAPPLMDLGDAAWLEPMFKHSGGL	180
DB	121	KOLYGSLNKKCLGLKEVARVSPFHGFTYCGPDEAPPLMDLGDAAWLEPMFKHSGGL	180
QY	181	ELVGPGRKVIVKANKKAPAEFVGDAYHVGMTTHASSLSRSGESIFSSLAGNMAALPEAGAL	240
DB	181	ELVGPGRKVIVKANKKAPAEFVGDAYHVGMTTHASSLSRSGESIFSSLAGNMAALPEAGAL	240
QY	241	QMTSKYSGMGVLMGYSVSHADLVPELMAFGAKQKQELNKEIGDVRIRYRSHLNCIV	300
DB	241	QMTSKYSGMGVLMGYSVSHADLVPELMAFGAKQKQELNKEIGDVRIRYRSHLNCIV	300
QY	301	PNNSNMLTCSGVFKWNPIDANTTEWTAIYEKMPEDLKRRLADSVORTGPGPFMS	360
DB	301	PNNSNMLTCSGVFKWNPIDANTTEWTAIYEKMPEDLKRRLADSVORTGPGPFMS	360
QY	361	DDNDNMTASQNGKKYQSRSDLSNLGFGEDDYGDAYPGVYGSAIGETYSRGFAY	420
DB	361	DDNDNMTASQNGKKYQSRSDLSNLGFGEDDYGDAYPGVYGSAIGETYSRGFAY	420

Db 361 DDNDNMFASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNMAEFERHASTTWHTELKTDDR 449  
Db 421 QAHVSSNMAEFERHASTTWHTELKTDDR 449

## RESULT 2

S27632

naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large chain

C:Species: Pseudomonas sp.

C:Date: 06-Jan-1995 #sequence-revision 06-Jan-1995 #text-change 15-Oct-1999

C:Accession: S27632; B49343

R:Denome, S.A.; Young, K.D.

Submitted to the EMBL Data Library, February 1992

A:Description: Cloning and molecular characterization of genes involved in metabolism of

A:Reference number: S27631

A:Accession: S27632

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-462 &lt;DEN&gt;

A:Cross-references: EMBL:M60405

R:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.

J. Bacteriol. 175, 6890-6901, 1993

A:Title: Metabolism of dibenzothioophene and naphthalene in Pseudomonas strains: complete

A:Reference number: A49343; MUID:94042852; PMID:8226631

A:Accession: B49343

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 14-462 &lt;DE2&gt;

A:Cross-references: GB:M60405; NID:g151195; PIDN:AAA16125.1; PID:g294351

A:Experimental source: strain C18

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [2Fe

C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein

F:84-132/Domain: Rieske [2Fe-2S] homology &lt;RSK&gt;

F:94,96,114,117/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 99.8%; Score 2404; DB 2; Length 462;

Best local similarity 99.8%; Pred. No. 1.2e-183;

Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNNKILVSSGSLQKHLHGDEELFOHELKTIFARNWLEFLTHDSLIPAGDYVYAKMG 60  
Db 14 MNYNNKILVSSGSLQKHLHGDEELFOHELKTIFARNWLEFLTHDSLIPAGDYVYAKMG 73  
QY 61 IDEVIVSRQDGSIRAFILNVCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEE 120  
Db 74 IDEVIVSRQDGSIRAFILNVCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEE 133  
QY 121 KDLYGESLNKKCLGLKEVARVESFHGFLYGCFOEDAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 134 KDLYGESLNKKCLGLKEVARVESFHGFLYGCFOEDAPPLMDYLGDAAWYLEPMFKHSGGL 193  
QY 181 ELVGPGRKVVYIKAMWKAPEENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPRPEGAGL 240  
Db 194 ELVGPGRKVVYIKAMWKAPEENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPRPEGAGL 253  
QY 241 QMTSKYSGMGVLMDGYSGVYASADLPVELMAFGAKOERLKEIGDVARIRYRSHLNTCV 300  
Db 254 QMTSKYSGMGVLMDGYSGVYASADLPVELMAFGAKOERLKEIGDVARIRYRSHLNTCV 313  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVORTLGPAGFWES 360  
Db 314 FPNNSMLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVORTLGPAGFWES 373  
QY 361 DDNDNMFASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 374 DDNDNMFASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 433  
QY 421 QAHVSSNMAEFERHASTTWHTELKTDDR 449  
Db 434 QAHVSSNMAEFERHASTTWHTELKTDDR 462

## RESULT 3

JN0644

naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large cha

N:Alternate names: nahac protein

C:Species: Pseudomonas putida

C:Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 20-Aug-1999

C:Accession: JN0644

R:Simon, M.J.; Osslund, T.D.; Saunders, R.; Ensley, B.D.; Suggs, S.; Harcourt, A.; Su

Gene 127, 31-37, 1993

A:Title: Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida st

A:Reference number: JN0640; MUID:93252277; PMID:8486285

A:Accession: JN0644

A:Molecule type: DNA

A:Residues: 1-449 &lt;STM&gt;

A:Cross-references: GB:M83949; NID:g151384; PIDN:AAA25902.1; PID:g151387

C:Comment: This protein is a member of naphthalene dioxygenase multicomponent enzyme

C:Genetics:

A:Gene: nahac

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [

C:Keywords: 2Fe-2S; heterotetramer; metalloprotein; oxidoreductase; Rieske iron-sulfu

F:71-119/Domain: Rieske [2Fe-2S] homology &lt;RSK&gt;

F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 97.4%; Score 2346; DB 2; Length 449;

Best local similarity 96.7%; Pred. No. 4.8e-179;

Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

QY 1 MNYNNKILVSSGSLQKHLHGDEELFOHELKTIFARNWLEFLTHDSLIPAGDYVYAKMG 60  
Db 1 MNYNNKILVSSGSLQKHLHGDEELFOHELKTIFARNWLEFLTHDSLIPAGDYVYAKMG 60  
QY 61 IDEVIVSRQDGSIRAFILNVCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEE 120  
Db 61 IDEVIVSRQDGSIRAFILNVCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEE 120  
QY 121 KDLYGESLNKKCLGLKEVARVESFHGFLYGCFOEDAPPLMDYLGDAAWYLEPMFKHSGGL 180  
Db 121 KDLYGESLNKKCLGLKEVARVESFHGFLYGCFOEDAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPGRKVVYIKAMWKAPEENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPRPEGAGL 240  
Db 181 ELVGPGRKVVYIKAMWKAPEENFVGDAYHVGWTHASSLSRSGESIFSSLAGNAALPRPEGAGL 240  
QY 241 QMTSKYSGMGVLMDGYSGVYASADLPVELMAFGAKOERLKEIGDVARIRYRSHLNTCV 300  
Db 241 QMTSKYSGMGVLMDGYSGVYASADLPVELMAFGAKOERLKEIGDVARIRYRSHLNTCV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVORTLGPAGFWES 360  
Db 301 FPNNSMLTCSGVFKVWNPIDANTTEVWTYAIVEKMPEDLKRRLADSVORTLGPAGFWES 360  
QY 361 DDNDNMFASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 361 DDNDNMFASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSNMAEFERHASTTWHTELKTDDR 449  
Db 421 QAHVSSNMAEFERHASTTWHTELKTDDR 449

## RESULT 4

C55217

polycyclic aromatic hydrocarbon dioxygenase (EC 1.14.12.-) iron-sulfur protein large

C:Species: Pseudomonas putida

C:Date: 05-May-1995 #sequence-revision 05-May-1995 #text-change 20-Jun-2000

C:Accession: C55217

R:Takizawa, N.; Kaida, N.; Toriogo, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara

J. Bacteriol. 176, 2444-2449, 1994

A:Title: Identification and characterization of genes encoding polycyclic aromatic hy

A:Reference number: A55217; MUID:94209249; PMID:8157615

A:Accession: C55217

A:Status: preliminary

A: Molecule type: DNA  
A: Residues: 1-449 <TAK>  
A: Cross-references: GB:AB040059; GB:D16629; NID:g2189972; PID:BA020391.1; PID:g391044  
C: Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-2S]  
C: Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F: 71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F: 81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status predicted

Query Match 95.2%; Score 2292; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 9,7e-175;  
Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNYNKIIVSSGSLQKLLHIGDEELFOHELKTIFARWMLFTHDSLIPAPGDVYTKMG 60  
Db 1 MNKKIKIVSSGSLQKLLHIGDEELFOHELKTIFARWMLFTHDSLIPAPGDVYTKMG 60  
QY 61 IDEVIVSRNDGSIKRAFLNVCRRHNGKTLVSVEAGNAKGFVSYHMGFGSGNGELQVPFE 120  
Db 61 IDEVIVSRNDGSIKRAFLNVCRRHNGKTLVSVEAGNAKGFVSYHMGFGSGNGELQVPFE 120  
QY 121 KDLYGESLNKKCLGLKEVARYESFHGFTYGCDFDQAPPLMYLGDAAVYLEPMRHSGCL 180  
Db 121 KELVGESLNKKCLGLKEVARYESFHGFTYACIDQAPPLMYLGDAAVYLEPIRHSGL 180  
QY 181 ELVGPGRKVVYIKAMKAPAEVFGDVAHYGWTTHASSLSGSEIFSSLSGNALPPEGAGL 240  
Db 181 ELVGPGRKVVYIKAMKAPAEVFGDVAHYGWTTHASSLSGSEIFSSLSGNALPPEGAGL 240  
QY 241 QMTSKYGGMGVLMDDGYSVGHSHADLPPELMAFGAKOERLNKETIGDVARIRSHLNTCV 300  
Db 241 QMTSKYGGMGVLMDDGYSVGHSHADLPPELMAFGAKOERLNKETIGDVARIRSHLNTCV 300  
QY 301 FPNNSMLTCGSGVFKWNPIDANTTEVWTVAIVKDMPEDLRLADSVQRLGPAEFMES 360  
Db 301 FPNNSMLTCGSGVFKWNPIDANTTEVWTVAIVKDMPEDLRLADSVQRLGPAEFMES 360  
QY 361 DDNDMMETASONGKRRKYOSRDSLNLGFGEDVYGDVAIPGVGKSAIGETSYRGFYAY 420  
Db 361 DDNDMMETASONGKRRKYOSRDSLNLGFGEDVYGDVAIPGVGKSAIGETSYRGFYAY 420  
QY 421 QAHVSSNMWAEFEHASTTWHTELTFTDR 449  
Db 421 QAHVSSNMWAEFEHASTTWHTELTFTDR 449

RESULT 5  
JC5352  
2-nitrotoluene dioxygenase (EC 1.14.12.-) iron-sulfur protein large chain - Pseudomonas  
N: Alternate names: 2-nitrotoluene iron-sulfur protein alpha chain  
C: Species: Pseudomonas sp.  
C: Date: 27-May-1997 #sequence-revision 18-Jul-1997 #text-change 20-Aug-1999  
C: Accession: JC5352  
R: Parales, J.V.; Kumar, A.; Parales, R.E.; Gibson, D.T.  
A: Title: Cloning and sequencing of the genes encoding 2-nitrotoluene dioxygenase from *Pseudomonas*  
A: Reference number: JC5350; MUID:97128768; PMID:8973308  
A: Accession: JC5352  
A: Molecule type: DNA  
A: Residues: 1-447 <PAR>  
A: Cross-references: GB:U49504; NID:g173273; PID:AA040383.1; PID:g1773277  
A: Experimental source: strain JS42  
C: Comment: This enzyme catalyzes the addition of dioxygen to the aromatic nucleus of 2-nitrotoluene  
C: Genetics:  
A: Gene: nldAc  
C: Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-2S]  
C: Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F: 69-117/Domain: Rieske [2Fe-2S] homology <RSK>  
F: 79,81,99,102/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status predicted

Query Match 85.3%; Score 2054; DB 2; Length 447;  
Best Local Similarity 83.7%; Pred. No. 8.9e-156;  
Matches 376; Conservative 35; Mismatches 36; Indels 2; Gaps 1;

[illegible]

```

RESULT 6
T31134
naphthalene dioxygenase (EC 1.14.12.-) large chain - Sphingomonas aromaticivorans pJia
C:Species: Sphingomonas aromatiicivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 19-May-2000
C:Accession: T31134
R:Romane, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.
Submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom
A:Reference number: Z20992
A:Accession: T31134
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-459 <ROM>
A:Cross-references: EMBL.AF079317; NID:g3378261; PID:g3378275; PIDN.AAD03858.1
A:Genetics:
A:Gene: bphA1 .
A:Genome: plasmid pNL1
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein
F:80,82,100,103/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match          39.7%, Score 957; DB 2; Length 459;
Best Local Similarity 43.9%; Pred. No. 2.5e+68;
Matches 192; Conservative 71; Mismatches 162; Indels 12; Gaps 7;

QY      8 LVSEGLSQKHLIHGDELFQHELKTI FARNMFLPTHSLIPAPEDVYTAKMGIDEIVYS 67
        || : || : |:::| || :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db       7 LVDMNASQSRREVDPODVYOLELERIFSCWLMIGHDSLVPPKPDFTITTYMAEDRYLLS 66
        || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY      68 RONDSTRIRALFWCHRRGKTLVSVEAGNAKGVCYSHGMSGNSCELOSVPFEKDYGES 127
        || ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db       67 ROPDSELAFINISCTHRGNQICHADSSAKAFVCNTHWGFGGDSLDVYMEEKCYHSD 126
        || ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY    128 LKKKCLGKEVARVESFGFIYCFCDFQEAPLIMDLIGAAAYLEPMFK-HSGGLETVGP 186
        || || : || : || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db    127 LDKSKLGIAPLI-RVETTYGFIICGHDPAPSLIEDYLGSFYCIYLDIIMGPRPGELLCP 185
        || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

QY    187 GKVIANKWKAPENVGDADYHVGMTTHASSLRGSEIFSSLAGNAA-LPPEGAGLQMTSK 245
        || :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

```

Db 186 LKSTLACNMKVPTEBNFYDGYHVGWTHAALOMIGELAGLSGNRADMPFDLLOFTMR 245  
QY 246 YGSGMGLMDGYSYVH--SADLPELMAFGAKOERLKEIGDVARIRYRSHLNTCTVPPN 303  
Db 246 HGHGFGGLIDNAATIAIHVKRQGYVYLETETRGIRKREKPE---RERLYVGHMWTSLPPN 301  
QY 304 NSMLTCSGVFWNPIDANTTEVWTAIVEKDMPEDLKRRRLADSVQRTLPAGFEMESDNN 363  
Db 302 CSFLYGTNTRFKIMHPRGHELEVWYTMVPRKNADTETKRSLQREAIRSFGAGLTESDDG 361  
QY 364 DNNETASONGKKYOSRSDLSNCFGEDVYG-DAVYPGVYKSAIGETSTRGTFRAYQA 422  
Db 362 ENMSATAYNNNGITTRKGRM--NSSMGKDRGPPHPYPGIIVGSPFISGTYSRGYFRWQE 419  
QY 423 HVSSNAEFEEHASTW 439  
Db 420 MLDAPDMAAIRANDTW 436

## RESULT 7

T31256  
terminal oxygenase component large chain homolog - Sphingomonas aromaticivorans plasmid  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31256  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A:Reference number: Z20992  
A:Accession: T31256  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-450 <ROM>  
A:Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378397; PIDN:AD03980.1  
C:Genetics:  
A:Genome: Plasmid pNL1  
A:Note: bphala  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

## Query Match

Best Local Similarity 32.18; Score 774; DB 2; Length 450;  
Matches 157; Conservative 88; Mismatches 184; Indels 6; Gaps 6;

QY 1 MNTNKKLVYSSGSLQKHLHGDELEFOHELKTIFARNMLFETHDSLIPAPGDYVTAKMG 60  
Db 1 MNDSIADLVBSRGRQSRISYASEDIYRQELERIFGRCLVLTHTSOILPKRGEDYRTFMG 60  
QY 61 IDEYIVSRONDGSTRAEFLVNCRHGRKTLVSEAGNAKGFVSYHGWGFGSNGELOSVPE 120  
Db 61 EDVYIVRQKDSIKAFELNSCTHGRNCRADRGNAFAFTCNHGMSPSPGALSGVPLE 120  
QY 121 KDLYGSELNKKCLGLKEVARYESHFGFIYCGFDOEAPRLMYLGDAAVTLER-MKKHSGG 179  
Db 121 NEAYFGEIDRTRKFLIPVTKYAEYKGLVFGCDANSPDLDTLGDAKFFLDVWLDAMPGG 180  
QY 180 LEIVGPRGVYIKAKMKAPAEENFYGDAYHVGWTHASSLSRSESI-FSSLA-GNALPREPG 237  
Db 181 SALLGETQKAVLGTNNKLPVENVCGDGYHLCMAHAGMAAQAOSMDLTLSLNGNSGVDDIG 240  
QY 238 AGLOMTSKYSGMGLVLDGYSYHSHADLVPELMAFGAKOERLKEIGDVAR-ITYRSHL 296  
Db 241 -GSLVAGMGNHMLVALDGVSYAFYPPDKPILEYLEARNRQTVIDRLGEVGRQWGAQV 299  
QY 297 NCYVFPNNKSLTCSGVYKVNPNPIDANTTEVWTAIVEKDMPEDLKRLADSVQRTLPAG 356  
Db 300 NITIFPMIQLPLGLNMYHPKPGPGOIEQWTMAEMDMPEAVAKAQLLENQCLTFGAG 359  
QY 357 FWSDDNDNNMETASONGKKYOSRSDLSNCFGEDVYGDAVYPGVYKSAIGETSYGFG 416  
Db 360 LFDNDGDNLTACTESGRGWTQAQMDVYTNMALRSGKREG-FPDIAAGLVASEHNOXYF 418

QY 417 YRAYQAHVSSSNMAE 431  
Db 419 YRRQEHMMAETMAE 433

## RESULT 8

A65031  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escheri  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: A65031; S49292  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shaoh, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: A65031  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <BLAT>  
A:Cross-references: GB:AE000340; GB:U00096; NID:g1788883; PIDN:AACT5591.1; PID:g17888  
A:Experimental source: strain K-12, substrain MG1655  
R:Tullin, E.; Gasser, F.; Biville, F.  
submitted to the EMBL Data Library, September 1994  
A:Description: Cloning and sequencing of an E. coli gene homologous to dioxygenase of  
A:Reference number: S49292  
A:Accession: S49292  
A:Molecule type: DNA  
A:Residues: 1-19, 'A', 21-383, 'ATAATATACWKK', 397-398, 'A51', 'R', 'A53', 'SAAATTAATLTISFQKIPL  
A:Cross-references: EMBL:Z37966; NID:g350595; PIDN:CAAB6018.1; PID:g350596  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:75-123/Domain: Rieske [2Fe-2S] homology <RSK>  
F:85,87,105,108/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 30.68; Score 737; DB 2; Length 453;  
Best Local Similarity 36.48; Pred. No. 8.3e-51;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGDELEFOHELKTIFARNMLFETHDSLIPAPGDYVYAKKGIVGVYVSRONDGSTRAEFLN 79  
Db 24 IYTPDITQLERLIFGRCLVLTHTSOILPKRGEDYRTFMG 83  
QY 80 VCRHGRKTLVSEAGNAKGFVSYHGWGFGSNGELOSVPEKDLGSLNKKCLGLKEVA 139  
Db 84 OCRHARMVSYADCGNPRAFPCPYHGMSTGNCBLIDVPRAYPGGLCKSHMGLNEVP 143  
QY 140 RVESHGFIYCGFDOEAPRLMYLGDAAVTLERPMF-KHSGLELVGPRGVYIKAKMKAP 198  
Db 144 CVESYKGLIFGNMPTSAFLRDYLDIAWLDGMLDRREGTEIVGVQKVVINCWKKFP 203  
QY 199 AENFYGDAYHVGWTHASSLSRSESI-FSSLA-GNALPREPG 234  
Db 204 AEQFASDQYHALLFHSASHAVOYLGAKKDGSCKRLGD-----GOTAPRVETAKDALQEG 256  
QY 235 --PEGAGLOMTSKYSGMGLVLDGYSYHSHADLVPELMAFGAKOERLKEIGDVARIRY 292  
Db 257 QDGHOSGFEFFKEKPDANVW--DCGAVSYRETYAE-----ABGRLEGVRLRL 303  
QY 293 RSHLNTCYVFPNNKSLTCSGVYKVNPNPIDANTTEVWTAIVEKDMPEDLKRLADSVQRTL 352  
Db 304 AGHNN--TFPTLSWINGATLRLVWHPRGPDQVEWMAFCIDDKAASDEYKAFENSATRAP 361  
QY 353 GRAGFWDDNDNNMETASONGKKYOSRSDLSNCFGEDVYGDAVYPGVYKSAIGETS 412  
Db 362 GPAGFLDQDSENNCEIOKLKGRANRSLKCLEMGVGOEKRRDDGIPGILT-NYIFSETA 420  
QY 413 YRGFYRAYQAHVSSSNMAEETWHTLT 445  
Db 421 ARGWQRRADLSSSESNQEVLDKTAAYQOEVYK 453

## RESULT 9

D91054 biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 31-Dec-2001  
C:Accession: D91054  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gas/water, N.; Yasunaga, T.; Kuhnara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome  
A:Reference number: M01D:21156231; PMID:11558796  
A:Accession: D91054  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-453 <HAV>  
A:Cross-references: GB:BA000007; PIDN:BA836827.1; PID:q13362875; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
C:Genetics:  
A:Gene: ECS3404  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-  
:Keywords: oxidoreductase

Query Match	30.6%	Score 7.37	DB 2	Length 453
Best Local Similarity	36.4%	Pred. No. 8,3e-51		
Matches 165	Conservative 64	Mismatches 174	Indels 50	Gaps 8
QY	20	IGHDEELFOHELKTIFARNMLFLTHDSLIPAPGDYVAKMGIDEIVYSRQDGSIRAFLN	79	
DB	24	IYTPDPIYQLELERLFGRCWFLFLAHESQIPKPGDFPNTYMGEDAVVYVRQKDGSIKAFN	83	
QY	80	VCRRGKTIYVEVGNMKGFCYSYHGWGFGSNGELQSPFEKDLGSELNKKCLGIKFEYA	139	
DB	84	QCRHRAARVSTADGQNTAFPCPHGWSYGINGELIDVPLPRAPQGLCKSHWGLENVP	143	
QY	140	RVESEHGEIYSCFQDEAPRLMDYLGDAAWYLEPMF-KHSGGLELVPGKVVIRANKAP	198	
DB	144	CVESEYKGLIFENMPTSPAGLRDYLGDIAWYLDGMIDRREGGTEIYGVQWVINCWKFP	203	
QY	199	AENPVDAHYHGWTHASL-----RSGESIFSSLAGNAALP-----	234	
DB	204	AEQRASDQYALHEFHASVVOYLKADGSDKRLGD-----GOTARVWETAKDALQFG	256	
QY	235	--PEGAGLQMTSKYSGGNGVLMDYSGVHASADVLPELMAFGAKOERLNKEIGDVRARIY	292	
DB	257	QDGGSGCGFEFEKPDANWV--DGAVSSTYRETYAE-----AEQRLGEVRLRL	303	
QY	293	RSHTNCTYFEPNNMVLTCOSGVFKVYNNPIDANTTEVWTYAYVEKMDPEDKRLABSVQTL	352	
DB	304	AGHNN--LEPFLSLWNLGATLRVWHPRGPDQVEWAFETITDKASDEYKAFENSATPAF	361	
QY	353	GPAGFWESDDNNNETASQNKKKYQSRSDSLSLNGFEIDYGGDAVYGVGVKSAIGETS	412	
DB	362	GPAGLEDDQDSENNCEIQKLKLGHRARNSKCLFEMGLGQGERRDDGJIGIT-NYIFSETA	420	
QY	413	YRGFYRAVQAIVSSNNAEEFHASSTWITTELTK	445	
DB	421	ARGMIGQRWADLLSESWQEVLDKTPAAVQOEYVK	453	

```

RESULT 10
C85898
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escherichia coli
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Dec-2001
C:Accession: G85898
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.;
Hiller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.;
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85898
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-453 <STO>

```

A:Cross-references: GB:AE005174; NID:g12516944; PDB:1AAG57651.1; GSDDB:GN00145; UMG:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: hcaal  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [hcaal]  
C:Keywords: oxidoreductase

Query Match	30.6%;	Score 737;	DB 2;	Length 453;
Best Local Similarity	36.4%;	Pred. No. 8.3e-51;		
Matches 165;	Conservative 64;	Mismatches 174;	Indels 50;	Gaps 8.

[illegible]

```

RESULT 11
T31258
aromatic oxygenase large chain - Sphingomonas aromaticivorans plasmid pNL1
C:Species: Sphingomonas aromaticivorans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 19-May-2000
C:Accession: T31258
R:Romaine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.
Submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom
A:Reference number: Z20992
A:Accession: T31258
A:Status: preliminary: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1:455 <ROM>
A:Cross-references: EMBL:AF079317; NID:g33378261; PID:g3378399; PIDN:AD03982.1
C:Genetics:
A:Genome: plasmid pNL1
A:Note: bphA1b
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein
F:88,90,108,111/matching site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status
Query Match 30.1%; Score 725.5; DB 2; Length 455;
Best Local Similarity 36.1%; Pred. No. 6.9e-50;
Matches 153; Conservative 83; Mismatches 169; Indels 19; Gaps 8;

Cy 2 NYNNKIIVSESGSKHLIHGDELFQHELTIFARKMLFLTHSLILAPGDYIARKGI 61
11 DYSRMDLKEGMLNR--IFSDADIVEEELIFARSWLFAVHESQIPSSGPFITHMGE 68

```

QY	62	DEIVSNOQNGSIAFLVNCRRHRTKIVSVEAGNAKGFCSYHNGMGRGSGELOSVEFEK	121
QY	62	DEIVSNOQNGSIAFLVNCRRHRTKIVSVEAGNAKGFCSYHNGMGRGSGELOSVEFEK	121
Db	69	DAIVIAQOPQSGISIVMLNLSCPBGRNKKVCFADAGNRRFEVNCYHGMAFD7AGDGMHEEY	128
QY	122	DLVGESLNKKCLGJEKVARVESFHFICYGCFQOEAAPLMDYLIDDAMYLEPME-KHSGGL	180
Db	129	CYIAGDIDLFKNHGLKNAKAVGNKYGLVFATFNSDAPSLSEMLDGFRIYLLMILDNEGGT	188
QY	181	ELVGPPEKVVYIKANWKAPAEENFVGDYIHWMTHASSLRSGESIFSSLAGNAALPPEGAGL	240
Db	189	EFIGGCIKISYISANMKFCVENFIDGAYHAGWTHDSTR-----SMNNGQPFP-----I	237
QY	241	QMTSKYSGSM-GVLMD-GYSGVHSDLV--PELMAFGAKOEPLNKETICGYRARIYSHL	296
Db	238	DMONSYHASVNGHMEFEETEGVDLELLGRPKVMDYNNKIRPKMAELRSGMRSKIFGSA	297
QY	297	NCTVFPNNSMLTCSGVFPKVMPIDANTTEVWYAIYEKMPEDLKRRLADSVORTLGPAG	356
Db	298	SASIFPNAVSLPGISTIRQOMQPKGPQOFELKTWYIANKNMPDUIKEEVYKGVWQTFECPGG	357
QY	357	FWESDNDNNHETASQNGKKYOSRDSLLSNLGFGEVDYDAPPGVVGKSAIGTSTYRGF	416
Db	358	TFEMDGEENNENCTIVNNGVYTRHERLHYRCGIGRQIEHDTL-PGLIYVYRQYDANRGF	416
QY	417	YRAY 420	
Db	417	YORW 420	

RESULT 12  
JN0812  
benzene 1,2-dioxygenase (EC 1.14.12.3) terminal oxygenase component large chain - Pseudomonas putida  
C:Species: Pseudomonas putida  
C:Date: 19-May-1994 #sequence-revision 06-Jan-1995 #text-change 21-Jul-2000

C:Accession: JN0812  
R:Tan, H.M.; Tang, H.Y.; Joannou, C.L.; Abdel-Wahab, N.H.; Mason, J.R.  
Gene 130, 33-39, 1993  
A:Title: The *Pseudomonas putida* M22 plasmid-encoded genes for benzene dioxygenase are un-  
A:Reference number: JN0810; M01D:95345820; PMID:5344526

A:Accession: JN0812  
A:Molecule type: DNA  
A:Residue: 1-450 <TRAN>  
A:Cross-references: GB:U04642; GB:U04643; NID:AA17758.1; PID:G309855  
A:Experimental source: strain ML2  
A:Comment: This enzyme is involved in catalyzing the oxidation of benzene to cis-1,2-dihydro-  
A:Comment: The reduced enzyme catalyzes the oxidation of benzene to cis-benzene dihydrodiol  
A:Genetics:

A:Gene: bedc1  
 A:Genome: plasmid  
 C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-2S; aromatic hydrocarbon catabolism; metalloprotein; oxidoreductase; Rieske-134/Domain: Rieske (2Fe-2S) homology <RSC>  
 F:86,98,116,119/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status profl

Query match	27.18; Score 652.5; DB 1; Length 450;
-------------	---------------------------------------

Best Local Similarity 34.68; Pred. No. 4.4e-44;  
Matches 148; Conservative 69; Mismatches 174; Indels 37; Gaps 11;

**Qy**    20 IHCDEL FQHLEKTI FARNWLELTHDSLIPAGDYVTAAMGIDEIVSRONGSIRAFEN 79  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Dd**    35 IYDDELLYQLELEVERFARSWLLIGHETHIRKPGRDYFTTYMGEDPVVVVRKDKDSIAVEN 94

**Qy**      80 VCRHRKTLVSEVGNAKGFVCYSYHGMGFGSNGELQSVPFEKDLYGESLNKKCLGLE-- 137  
      |||||+|||||+|||||+:||:|||||:|                  || ||  
**Db**      95 QCRHRRGMRICRSPAGNKAFFCSYHGNAVDPDAGNLINLVPEAESFA-----CLDKREWS 148

QY 138 --VARYESFHGTYGCFDQEARPLMDYLGDAWYLEPME-KHSGGLELYGPPGKVIKAN 194  
| | | | : : : : | | | | : : : : | : : : | | |  
149 PLKARFETVYGTIFRANNDENATIDYGVYCPAEVEWNUYIDDEPNAEGREVTCTCKRMVTCN 208  
150

QY WKPAENFVGDAYHVGT-HASSLRGESIFSSLAGMALLPREGAGLOMTSKYSGMGVYL 253

[illegible]

RESULT 13  
A36516  
Toluene dioxygenase (EC 1.14.12.11) terminal oxygenase component large chain - Pseudomonas putida C1  
C:Species: Pseudomonas putida  
C:Dates: 15-Feb-1991 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
C:Accession: A36516  
R:Zylstra, G.J.; Gibson, D.T  
J. Biol. Chem. 264, 14940-14946, 1989  
A:Title: Toluene degradation by Pseudomonas putida F1. Nucleotide sequence of the toluene  
A:Reference number: A36516; MUID:89359301; PMID:2670929

A:Accession: A36516  
A:Molecule type: DNA  
A:Residues: 1-450 <ZYL>  
A:Cross-references: GB:J04996; NID:g151600; PIDN:AAA26005.1; PID:g151601  
A:Experimental source: strain F1

C:Genetics: 1  
A:Gene: t00c1  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain: Rieske [F:86-134/Domain: Rieske (2Fe-2S) homology <RSK> F:96,98,116,119/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 27.08; Score 650.5; DB 1; Length 450;

Best Local Similarity	34.48;	Pred: No. 6.4e-44;							
Matches	145;	Conservative	71;	Mismatches	181;	Indels	25;	Gaps	9;

**OY**      20 IHGDELLFOHEKTFEARNWLELTHDSLLPBGDVTAKMGIDEIVSRÖNDGSTRAFLN 79  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
**Dd**      35 IYTDEDLIQLEIFERFAASWMLIGHETOIRKPBGDIITYMGGEDPPVVVRÖMDASTAFLN 94

QY 80 VCRHKGKTLVSEAGNMGKGFVCSYHGKMGSGNGELQSVPEEKDYLGSLLKCKLGF-- 137  
 |||| : :||||| ||||| : : |||| : : || ||  
 Db 95 QCRHGMGRICRADACNAKFTCSYHGMAVDPRAGNLVNNPYELESFA-----CLKKKMS 148

Qy 138 --VARVESEHGFITCYCFQOEAPPLMDYLGDAWYLEPMF-KHSGLELVGPBGKVIKAN 194  
| | | | | : : : | | | | : : : | : : | : | |  
Db 149 PLKAVEVEYKGLIFENWBNENAVDLDTYLGAEKFFYMDHLDRIEAGTEAIPGVOKWVPCN 208  
| | | | | : : : | | | | : : : | : : | : | |

OY    195 WKPAENFVGDAYHYGWT--HASSLSRG--ESIFSSLAGNNAALPEEGAGLQMTSKYGSGMG 251  
      || || | || | || | : : : :  
 Db    209 WKAFAEFGCSMDVHAAGTSHSGLIAGLPEDL-----EMADLAPPTVKOVYBASGCHGS 263

25	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--

[illegible]

D0	3 2 2	P O I N T I V K T W H P R G F N E V E W M A F I V V D A D A P D D I K E E F R K O I L K I E F S A G G V E E D D G E N W	3 8 1
Qy	3 6 8	T A S O N G K K Y O S R D D L L N L G F G E D V G D A Y Y P G V W G S A I G E M S Y R G F R A Y Q A H V S S	4 2 7







XX  
XX

PT Novel naphthalene dioxygenase mutant having a specific amino acid



Db 121 KDLYESLKKKCLGLKEVARVESFHGFIYGCDFDQAPPLMDYLGDAAWYLEPMPFKHSGGL 180  
 QY 181 ELVGPGRKVIKANKKAPAEENVGDAYHVGMTTHASSLSRGSIFSSLAGNAPLPEAGL 240  
 Db 181 ELVGPGRKVIKANKKAPAEENVGDAYHVGMTTHASSLSRGSIFSSLAGNAPLPEAGL 240  
 QY 241 QMTSKYSGMGVLMGYSVSHSDLYPELMARFGAKQERLKEIGDVRARIRSHLNCIV 300  
 Db 241 QMTSKYSGMGVLMGYSVSHSDLYPELMARFGAKQERLKEIGDVRARIRSHLNCIV 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWTAIYEKMPEDLKRRLADSVORTLGPAGFWS 360  
 Db 301 FPNNSMLTCSGVFKVWNPIDANTTEWTAIYEKMPEDLKRRLADSVORTLGPAGFWS 360  
 QY 361 DDNDNMETASQNGKKYQSRDLSNLGFGEDVDYDAVYPGVGSAIGETSYRGFYRAY 420  
 Db 361 DDNDNMETASQNGKKYQSRDLSNLGFGEDVDYDAVYPGVGSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449  
 Db 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449

RESULT 3  
 AAB12565  
 ID AAB12565 standard; Protein: 449 AA.  
 XX AAB12565;  
 AC  
 DT 09-NOV-2000 (first entry)  
 DE Naphthalene dioxygenase mutant F352V protein sequence SEQ ID NO:2.  
 XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 XX Pseudomonas sp.  
 OS Synthetic.  
 OS  
 PN WO200037480-A1.  
 PD 29-JUN-2000.  
 XX 26-OCT-1999; 99WO-US25079.  
 PF 26-OCT-1999; 98US-0105575.  
 PR 26-OCT-1998; 98US-0105575.  
 XX (IOWA ) UNITV IOWA RES FOUND.  
 PA  
 PI Parales R, Gibson D, Resnick S, Lee K;  
 XX WPI: 2000-452174/39.  
 DR N-PSDB: AAA65339.  
 XX  
 PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT  
 XX  
 PS Claim 9; Page 58-60; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (-)- or (+)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which

CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

Sequence 449 AA:  
 SQ  
 Query Match 99.9%; Score 2405; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 3.1e-223;  
 Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYNKKILVSESGLSQKHILHGEDELFOHELKTIFARMWLFTHDSLIPAPGDVYATKMG 60  
 Db 1 MNYNKKILVSESGLSQKHILHGEDELFOHELKTIFARMWLFTHDSLIPAPGDVYATKMG 60  
 QY 61 IDEVIVSRQNDGSTRAPLNVCHRGKTLVSVAGNAKGVCSYHGMGSGNCELOSVPFE 120  
 Db 61 IDEVIVSRQNDGSTRAPLNVCHRGKTLVSVAGNAKGVCSYHGMGSGNCELOSVPFE 120  
 QY 121 KDLYESLKKKCLGLKEVARVESFHGFIYGCDFDQAPPLMDYLGDAAWYLEPMPFKHSGGL 180  
 Db 121 KDLYESLKKKCLGLKEVARVESFHGFIYGCDFDQAPPLMDYLGDAAWYLEPMPFKHSGGL 180  
 QY 181 ELVGPGRKVIKANKKAPAEENVGDAYHVGMTTHASSLSRGSIFSSLAGNAPLPEAGL 240  
 Db 181 ELVGPGRKVIKANKKAPAEENVGDAYHVGMTTHASSLSRGSIFSSLAGNAPLPEAGL 240  
 QY 241 QMTSKYSGMGVLMGYSVSHSDLYPELMARFGAKQERLKEIGDVRARIRSHLNCIV 300  
 Db 241 QMTSKYSGMGVLMGYSVSHSDLYPELMARFGAKQERLKEIGDVRARIRSHLNCIV 300  
 QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWTAIYEKMPEDLKRRLADSVORTLGPAGFWS 360  
 Db 301 FPNNSMLTCSGVFKVWNPIDANTTEWTAIYEKMPEDLKRRLADSVORTLGPAGFWS 360  
 QY 361 DDNDNMETASQNGKKYQSRDLSNLGFGEDVDYDAVYPGVGSAIGETSYRGFYRAY 420  
 Db 361 DDNDNMETASQNGKKYQSRDLSNLGFGEDVDYDAVYPGVGSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449  
 Db 421 QAHVSSSNMAEFEEHASSTWHTELTKTTDR 449

RESULT 4  
 AAB12566  
 ID AAB12566 standard; Protein: 449 AA.  
 XX AAB12566;  
 AC  
 DT 09-NOV-2000 (first entry)  
 DE NDO related complex alpha subunit protein sequence SEQ ID NO:14.  
 XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 XX Pseudomonas sp.  
 OS Synthetic.  
 OS  
 PN WO200037480-A1.  
 PD 29-JUN-2000.  
 PF 26-OCT-1999; 99WO-US25079.

XX 26-OCT-1998; 98US-0105575.  
 PR (IOWA ) UNIV IOWA RES FOUND.  
 XX  
 XX  
 PI Parales R, Gibson D, Resnick S, Lee K;  
 XX WPI: 2000-452174/39.  
 DR N-PSDB: AAA65340.

XX Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT

XX Claim 13; Page 99-100; 151pp; English.

XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidize an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA;

Query Match 99.9%; Score 2405; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 3.1e-223;  
 Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNTNKKLTVESGSLQKHLHGEDELFQHEKLTFFARNMFLTHDSLRPAQGVYTAAG 60  
 DB 1 MNTNKKLTVESGSLQKHLHGEDELFQHEKLTFFARNMFLTHDSLRPAQGVYTAAG 60  
 OY 61 IDEVTVSRQDGSIRAFLANVCRRHGKTLVSVEAGNAGFVSYGQWGSNGELQSVFE 120  
 DB 61 IDEVTVSRQDGSIRAFLANVCRRHGKTLVSVEAGNAGFVSYGQWGSNGELQSVFE 120  
 OY 121 KDLVGSINKKCLGLKEVARSEFHFGLYGCFOEAPLMDYLDGAAYLEPMFKHSGSL 180  
 DB 121 KDLVGSINKKCLGLKEVARSEFHFGLYGCFOEAPLMDYLDGAAYLEPMFKHSGSL 180  
 OY 181 ELVGPCKVYIKANKWKAPENFVGDAYHVGWTHASSLSGSEIFSSLAGNALPPEGAGL 240  
 DB 181 ELVGPCKVYIKANKWKAPENFVGDAYHVGWTHASSLSGSEIFSSLAGNALPPEGAGL 240  
 OY 241 QMSTKSGSGVGLMDGSGVSHADLPVLMFAGKQRLKEIGDVARYRSLNCTV 300  
 DB 241 QMSTKSGSGVGLMDGSGVSHADLPVLMFAGKQRLKEIGDVARYRSLNCTV 300  
 OY 301 FPNNSMLTSGVFKVNPIDANTTEWYTAIVERKMPEDLKRRLADSVQRTLGPAGFWS 360  
 DB 301 FPNNSMLTSGVFKVNPIDANTTEWYTAIVERKMPEDLKRRLADSVQRTLGPAGFWS 360  
 OY 361 DDNDNMTASONGKRYOSRSDLSNLGFGEDYGDVAYPGVYKSAIGETSYRGFYRAY 420  
 DB 361 DDNDNMTASONGKRYOSRSDLSNLGFGEDYGDVAYPGVYKSAIGETSYRGFYRAY 420

OY 421 QAHVSSNMAEFHASTWHTLKTDDR 449  
 DB 421 QAHVSSNMAEFHASTWHTLKTDDR 449

RESULT 5  
 AAB12567  
 ID AAB12567 standard; Protein; 449 AA.

XX AAB12567;  
 XX 09-NOV-2000 (first entry)

DE NDO related complex alpha subunit protein sequence SEQ ID NO:15.

XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS Pseudomonas sp.  
 OS Synthetic.

PN WO200037480-A1.

PD 29-JUN-2000.

PF 26-OCT-1999; 99NO-US25079.

PR 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB: AAA65341.

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT

XX Claim 13; Page 100-102; 151pp; English.

XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidize an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

SO Sequence 449 AA;

Query Match 99.9%; Score 2405; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 3.1e-223;  
 Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYNKKILVSESGLSOKHLIHGDEELFQHELTIFARNMLFLTHDSLIPAPGDYATKMG 60  
 DB 1 MNYNKKILVSESGLSOKHLIHGDEELFQHELTIFARNMLFLTHDSLIPAPGDYATKMG 60  
 QY 61 IDEVIVSRONDGSIKRAFLVNCRRKGTILVSAGNAKGVCSYHGMGSGNELSVPE 120  
 DB 61 IDEVIVSRONDGSIKRAFLVNCRRKGTILVSAGNAKGVCSYHGMGSGNELSVPE 120  
 QY 121 KDLYGESLKKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGAAWYLEPMFKHSGSL 180  
 DB 121 KDLYGESLKKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGAAWYLEPMFKHSGSL 180  
 QY 181 ELVGPGRKVVIRANKKAPAEFVGDAYHYGWTTHASSLSRSGESIFSSLAGNAALPEPGAGL 240  
 DB 181 ELVGPGRKVVIRANKKAPAEFVGDAYHYGWTTHASSLSRSGESIFSSLAGNAALPEPGAGL 240  
 QY 241 QMTSKYSGMGVLMGYSVSHSADLVPPELMAFGAKOEFLNKIEIGDVARIRYRSHLNTCV 300  
 DB 241 QMTSKYSGMGVLMGYSVSHSADLVPPELMAFGAKOEFLNKIEIGDVARIRYRSHLNTCV 300  
 QY 301 FPNNSMLTCSGVYFKVWNPIDANTTEWWTYAIYEKDPEDLKRRLADSVQRTIGPAGFWES 360  
 DB 301 FPNNSMLTCSGVYFKVWNPIDANTTEWWTYAIYEKDPEDLKRRLADSVQRTIGPAGFWES 360  
 QY 361 DDNDNMEFASONGKRYQSRDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 DB 361 DDNDNMEFASONGKRYQSRDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449  
 DB 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449

RESULT 6  
 AAB12577  
 ID AAB12577 standard; Protein: 449 AA.  
 AC AAB12577;  
 XX  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE Naphthalene dioxygenase alpha subunit protein sequence SEQ ID NO:26.  
 XX  
 KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 XX polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 OS Pseudomonas sp.  
 PN WO200037480-A1.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 26-OCT-1999; 99WO-US25079.  
 XX  
 PR 26-OCT-1998; 98US-0105575.  
 XX  
 PA (IOWA) UNIV IOWA RES FOUND.  
 PI  
 FI Paraless R, Gibson D, Resnick S, Lee K;  
 DR WPI: 2000-452174/39.  
 DR N-PSDB; AAA65351.  
 XX  
 PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT  
 XX  
 PS Disclosure: Page 118-120; 151pp; English.  
 CC The present invention describes a naphthalene dioxygenase (NDO) or NDO

CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzofuran,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC naphthalene dioxygenase (NDO) from the Pseudomonas sp. strain NCIB  
 CC 9816-4, which is used in the exemplification of the present invention.

Sequence 449 AA:

Query Match 99.8%; Score 2404; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 3.9e-223;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNYNKKILVSESGLSOKHLIHGDEELFQHELTIFARNMLFLTHDSLIPAPGDYATKMG 60  
 DB 1 MNYNKKILVSESGLSOKHLIHGDEELFQHELTIFARNMLFLTHDSLIPAPGDYATKMG 60  
 QY 61 IDEVIVSRONDGSIKRAFLVNCRRKGTILVSAGNAKGVCSYHGMGSGNELSVPE 120  
 DB 61 IDEVIVSRONDGSIKRAFLVNCRRKGTILVSAGNAKGVCSYHGMGSGNELSVPE 120  
 QY 121 KDLYGESLKKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGAAWYLEPMFKHSGSL 180  
 DB 121 KDLYGESLKKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGAAWYLEPMFKHSGSL 180  
 QY 181 ELVGPGRKVVIRANKKAPAEFVGDAYHYGWTTHASSLSRSGESIFSSLAGNAALPEPGAGL 240  
 DB 181 ELVGPGRKVVIRANKKAPAEFVGDAYHYGWTTHASSLSRSGESIFSSLAGNAALPEPGAGL 240  
 QY 241 QMTSKYSGMGVLMGYSVSHSADLVPPELMAFGAKOEFLNKIEIGDVARIRYRSHLNTCV 300  
 DB 241 QMTSKYSGMGVLMGYSVSHSADLVPPELMAFGAKOEFLNKIEIGDVARIRYRSHLNTCV 300  
 QY 301 FPNNSMLTCSGVYFKVWNPIDANTTEWWTYAIYEKDPEDLKRRLADSVQRTIGPAGFWES 360  
 DB 301 FPNNSMLTCSGVYFKVWNPIDANTTEWWTYAIYEKDPEDLKRRLADSVQRTIGPAGFWES 360  
 QY 361 DDNDNMEFASONGKRYQSRDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 DB 361 DDNDNMEFASONGKRYQSRDLSNLGFGEDVYGDVAVPGVYKSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449  
 DB 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449

RESULT 7  
 AAB12579  
 ID AAB12579 standard; Protein: 449 AA.  
 AC AAB12579;  
 XX  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE Naphthalene dioxygenase mutant f352A protein sequence SEQ ID NO:33.  
 XX  
 KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;

polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS Pseudomonas sp.  
 OS Synthetic.  
 PN WO200037480-A1.  
 XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 XX 26-OCT-1998; 98US-0105575.  
 PR (IOWA ) UNIV IOWA RES FOUND.  
 PA Parales R, Gibson D, Resnick S, Lee K;  
 PI MPI: 2000-452174/39.  
 DR N-PSDB; AAA65353.  
 XX Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT

Claim 9; Page 128-129; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Sequence 449 AA:

Query Match 99.8%; Score 2403; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 4.9e-223;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 MNNNKLTVESGSGSOHLHGHDELFQHELTFFARWMLFTHDSLIPAPGDVYTKMG 60  
 1 MNNNKLTVESGSGSOHLHGHDELFQHELTFFARWMLFTHDSLIPAGDVTYTKMG 60  
 61 IDEVTSRQNDGSTRATLNCVCRHNGKTLVSEAGNAGFCVSGWGFSGNGELQSVFE 120  
 61 IDEVTSRQNDGSTRATLNCVCRHNGKTLVSEAGNAGFCVSGWGFSGNGELQSVFE 120  
 61 IDEVTSRQNDGSTRATLNCVCRHNGKTLVSEAGNAGFCVSGWGFSGNGELQSVFE 120  
 121 KDLVGSGLNKKCLGLKAVARVESHFGFTYGCFOEDAPPLMDYLGDAAVYLERPKHSGGL 180  
 121 KDLVGSGLNKKCLGLKAVARVESHFGFTYGCFOEDAPPLMDYLGDAAVYLERPKHSGGL 180  
 121 KDLVGSGLNKKCLGLKAVARVESHFGFTYGCFOEDAPPLMDYLGDAAVYLERPKHSGGL 180  
 181 ELVGPPEKVVYIKAMWKAPEVFGDAYHVGWTHASSLSRSGSSTSSLAGNAPREGAGL 240  
 181 ELVGPPEKVVYIKAMWKAPEVFGDAYHVGWTHASSLSRSGSSTSSLAGNAPREGAGL 240  
 241 QMNSKYSGMGVLMIDGYSGVHSADLVPDLMAFGAKQERLNKEIGDVARIRYRSHLNTCV 300

|||||

241 QMNSKYSGMGVLMIDGYSGVHSADLVPDLMAFGAKQERLNKEIGDVARIRYRSHLNTCV 300  
 301 FPNNSMLTSCGVEKVPWNPIDANTTEWTTAIVKMPEDIKRLADSVQRTLPAGFWS 360  
 301 FPNNSMLTSCGVEKVPWNPIDANTTEWTTAIVKMPEDIKRLADSVQRTLPAGFWS 360  
 361 DDNDNMTASONGCKKYSRSDLSNLGFGEDVYGDAVYPGVVGKSAIGETSYRGFRAY 420  
 361 DDNDNMTASONGCKKYSRSDLSNLGFGEDVYGDAVYPGVVGKSAIGETSYRGFRAY 420  
 421 QAHVSSNMAEFERASTWTELTCTDR 449  
 421 QAHVSSNMAEFERASTWTELTCTDR 449

RESULT 8  
 AAB12580  
 ID AAB12580 standard; Protein; 449 AA.  
 XX  
 AC AAB12580;  
 XX  
 DT 09-NOV-2000 (first entry)  
 XX  
 DE Naphthalene dioxygenase mutant F352T protein sequence SEQ ID NO:34.  
 XX  
 KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KW Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 XX  
 OS Pseudomonas sp.  
 OS Synthetic.  
 XX  
 PN WO200037480-A1.  
 XX  
 XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 XX 26-OCT-1998; 98US-0105575.  
 PR (IOWA ) UNIV IOWA RES FOUND.  
 PA Parales R, Gibson D, Resnick S, Lee K;  
 PI MPI: 2000-452174/39.  
 DR N-PSDB; AAA65354.  
 XX Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT

Claim 9; Page 129-131; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzothiophene, 9,10-dihydronaphthalene, or 9,10-dihydronaphthalene a corresponding dihydro dihydroxy compound. The polypeptides and the host



CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX

Sequence 449 AA:

Query Match 99.8%; Score 2403; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 4.9e-223;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVNNKILVSESGLSQKHLIHGDEELFQHEHKTIFARNMLFLTHDSLIAPGDIYAAKMG 60  
DB 1 MNVNNKILVSESGLSQKHLIHGDEELFQHEHKTIFARNMLFLTHDSLIAPGDIYAAKMG 60  
QY 61 IDEVIYSRONDGSIIRAFNLVNCRRHRTIVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIYSRONDGSIIRAFNLVNCRRHRTIVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPFGKVIYKANKKAPENFVGDAVHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
DB 181 ELVGPFGKVIYKANKKAPENFVGDAVHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVSHADYVPELMAFGAKOERLNKEIGVRRARIYRSHLNTCV 300  
DB 241 QMTSKYSGMGVLMDSYGVSHADYVPELMAFGAKOERLNKEIGVRRARIYRSHLNTCV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDPEDLKRRLADSVORTGPAFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDPEDLKRRLADSVORTGPAFWES 360  
QY 361 DDNDNMETASQNGKKYQSDLSNLGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSDLSNLGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFEEHASTWHTLTKTTDR 449  
DB 421 QAHVSSSNMAEFEEHASTWHTLTKTTDR 449

RESULT 9  
AAB12583  
ID AAB12583 standard; Protein; 449 AA.  
XX  
AC AAB12583;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352W protein sequence SEQ ID NO:58.  
XX  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
PI Parales R, Gibson D, Resnick S, Lee K.  
XX

DR WPI: 2000-452174/39.  
DR N-PSDB: AAA65376.  
XX

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT

Example 7; Page 142-144; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (1) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX

Sequence 449 AA:

Query Match 99.8%; Score 2402; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 6.1e-223;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVNNKILVSESGLSQKHLIHGDEELFQHEHKTIFARNMLFLTHDSLIAPGDIYAAKMG 60  
DB 1 MNVNNKILVSESGLSQKHLIHGDEELFQHEHKTIFARNMLFLTHDSLIAPGDIYAAKMG 60  
QY 61 IDEVIYSRONDGSIIRAFNLVNCRRHRTIVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
DB 61 IDEVIYSRONDGSIIRAFNLVNCRRHRTIVSVEAGNAKGFVCSYHGMFGSNGELQSVPE 120  
QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFKHSGGL 180  
QY 181 ELVGPFGKVIYKANKKAPENFVGDAVHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
DB 181 ELVGPFGKVIYKANKKAPENFVGDAVHYGWTTHASSLRSGESIFSSLAGNAALPPGAGL 240  
QY 241 QMTSKYSGMGVLMDSYGVSHADYVPELMAFGAKOERLNKEIGVRRARIYRSHLNTCV 300  
DB 241 QMTSKYSGMGVLMDSYGVSHADYVPELMAFGAKOERLNKEIGVRRARIYRSHLNTCV 300  
QY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDPEDLKRRLADSVORTGPAFWES 360  
DB 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIYEKDPEDLKRRLADSVORTGPAFWES 360  
QY 361 DDNDNMETASQNGKKYQSDLSNLGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
DB 361 DDNDNMETASQNGKKYQSDLSNLGFGEDYVGDVYPGVYKSAIGETSYRGFYRAY 420  
QY 421 QAHVSSSNMAEFEEHASTWHTLTKTTDR 449  
DB 421 QAHVSSSNMAEFEEHASTWHTLTKTTDR 449

RESULT 10



CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.7%; Score 2400; DB 21; Length 449;

Best Local Similarity 99.8%; Pred. No. 9.5e-223;

Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYATKMG 60  
 DB 1 MMYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYATKMG 60  
 QY 61 IDEVIVSRONDGSIKRAFLVNCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
 DB 61 IDEVIVSRONDGSIKRAFLVNCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
 QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDQAPLMDYLGDAAWYLEPMFKHSGGL 180  
 DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDQAPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPGRKVITANKKAPENFVGDAHYHGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
 DB 181 ELVGPGRKVITANKKAPENFVGDAHYHGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
 QY 241 QMTSKYSGMGVLMPOGYSVSHSADLPELMARFAGAKOEELNKEIGVRRIRYRSHLNCYV 300  
 DB 241 QMTSKYSGMGVLMPOGYSVSHSADLPELMARFAGAKOEELNKEIGVRRIRYRSHLNCYV 300  
 QY 301 FPNNMILTCSSGVFKVWNPIDANTTEVWYIAIVEKMDPEDLKRRLADSVORTLGPAGFWES 360  
 DB 301 FPNNMILTCSSGVFKVWNPIDANTTEVWYIAIVEKMDPEDLKRRLADSVORTLGPAGFWES 360  
 QY 361 DDNDNMETASQNGKRYQSODSLSLMGEGEDYGDVAPGVYKSAIETSTRYRAY 420  
 DB 361 DDNDNMETASQNGKRYQSODSLSLMGEGEDYGDVAPGVYKSAIETSTRYRAY 420  
 QY 421 QAHVSSSNMAEFEEHASTWTELTITTD 449  
 DB 421 QAHVSSSNMAEFEEHASTWTELTITTD 449

RESULT 12

AA012584

ID AA012584 standard: Protein: 448 AA.

XX AA012584;

DT 09-NOV-2000 (first entry)

DE Naphthalene dioxygenase mutant F352Y protein sequence SEQ ID NO:59.

KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;

KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;

KW polymer; resin; pharmaceutical; rubber industry; bioremediation.

XX Pseudomonas sp.

OS Synthetic.

PN WO200037480-A1.

XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 PF 26-OCT-1998; 98US-0105575.  
 PR (IOWA ) UNIV IOWA RES FOUND.  
 PA Parles R, Gibson D, Resnick S, Lee K;  
 DR WPI: 2000-452174/39.  
 DR N-PSDB: AAA65377.  
 PT Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 XX Example 7; Page 144-145; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydrophenanthrene or 3,4-dihydroxy-3,4-dihydrophenanthrene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydrophenanthrene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

SO Sequence 448 AA:

Query Match 99.6%; Score 2398; DB 21; Length 448;

Best Local Similarity 99.8%; Pred. No. 1.5e-222;

Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MMYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYATKMG 60  
 DB 1 MMYNNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFLTHDSLIPAGDYATKMG 60  
 QY 61 IDEVIVSRONDGSIKRAFLVNCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
 DB 61 IDEVIVSRONDGSIKRAFLVNCRRHKGTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPE 120  
 QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDQAPLMDYLGDAAWYLEPMFKHSGGL 180  
 DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQDQAPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPGRKVITANKKAPENFVGDAHYHGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
 DB 181 ELVGPGRKVITANKKAPENFVGDAHYHGWTHASSLRSGESIFSSLAGNAALPPGAGL 240  
 QY 241 QMTSKYSGMGVLMPOGYSVSHSADLPELMARFAGAKOEELNKEIGVRRIRYRSHLNCYV 300  
 DB 241 QMTSKYSGMGVLMPOGYSVSHSADLPELMARFAGAKOEELNKEIGVRRIRYRSHLNCYV 300  
 QY 301 FPNNMILTCSSGVFKVWNPIDANTTEVWYIAIVEKMDPEDLKRRLADSVORTLGPAGFWES 360  
 DB 301 FPNNMILTCSSGVFKVWNPIDANTTEVWYIAIVEKMDPEDLKRRLADSVORTLGPAGFWES 360

QY 361 DDNDNMTETASONGKRYOSRSDLSNLGFGEDVYGDVAVPGVGSALGETSYRGFYRAY 420  
 Db 361 DDNDNMTETASONGKRYOSRSDLSNLGFGEDVYGDVAVPGVGSALGETSYRGFYRAY 420

QY 421 QAHVSSSNMAEFHASTWHTELTKTTD 448  
 Db 421 QAHVSSSNMAEFHASTWHTELTKTTD 448

# RESULT 13

AAAB12569 standard; Protein: 449 AA.

AAAB12569;

09-NOV-2000 (first entry)

NDO related complex alpha subunit protein sequence SEQ ID NO:17.

Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO; inducible multi-component enzyme; alpha subunit; mutant; chiral diol; polymer; resin; pharmaceutical; rubber industry; bioremediation.

Pseudomonas sp. Synthetic.

WO200037480-A1.

29-JUN-2000.

26-OCT-1999; 99WO-US25079.

26-OCT-1998; 98US-0105575.

(IOWA ) UNIV IOWA RES FOUND.

Parales R, Gibson D, Resnick S, Lee K;

WPI: 2000-452174/39.

N-PSDB: AAA65343.

Novel naphthalene dioxygenase mutant having a specific amino acid substitution for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry and for carrying out bioremediation

Claim 13; Page 104-105; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzochlorophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Sequence 449 AA;

Query Match 99.5%; Score 2395; DB 21; Length 449;  
 Best Local Similarity 99.3%; Pred. No. 2, 9e-222;  
 Matches 446; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNVNNKLIVSESGLSQKHLHGDELFQHELTETFAARNMLFLHDSLIPAPGDVYAKMG 60  
 Db 1 MNVNNKLIVSESGLSQKHLHGDELFQHELTETFAARNMLFLHDSLIPAPGDVYAKMG 60  
 QY 61 IDEVIVSRQNDGSTRAPLNVCRHRRGKTLVSYEAGNAGFCVSYHGWGSGNGELQSVPE 120  
 Db 61 IDEVIVSRQNDGSTRAPLNVCRHRRGKTLVSYEAGNAGFCVSYHGWGSGNGELQSVPE 120  
 QY 121 KDLYGESLNRKCKLKEVAVRESFHGFYGCDFDEAPPLMDYLGDAAWYLEPMEKHSGL 180  
 Db 121 KDLYGESLNRKCKLKEVAVRESFHGFYGCDFDEAPPLMDYLGDAAWYLEPMEKHSGL 180  
 QY 181 ELVGPBGKVVYIKAMWKAPAEVNFVGDVAVHVGWTHASSLSRSGESTFSSLAGNAALPPEGAGL 240  
 Db 181 ELVGPBGKVVYIKAMWKAPAEVNFVGDVAVHVGWTHASSLSRSGESTFSSLAGNAALPPEGAGL 240  
 QY 241 QMTSKYSGMGVLDVGSVGSADLVPELMAFGAKQERLNKEIGDVRARATYSHLNCY 300  
 Db 241 QMTSKYSGMGVLDVGSVGSADLVPELMAFGAKQERLNKEIGDVRARATYSHLNCY 300  
 QY 301 FPNNSMLTFCSGVFWVWNPIDANTEVWTYALVEKDMPEDLKRLADSVORTLPAGFWES 360  
 Db 301 FPNNSMLTFCSGVFWVWNPIDANTEVWTYALVEKDMPEDLKRLADSVORTLPAGFWES 360  
 QY 361 DDNDNMTETASONGKRYOSRSDLSNLGFGEDVYGDVAVPGVGSALGETSYRGFYRAY 420  
 Db 361 DDNDNMTETASONGKRYOSRSDLSNLGFGEDVYGDVAVPGVGSALGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFHASTWHTELTKTTD 449  
 Db 421 QAHVSSSNMAEFHASTWHTELTKTTD 449

# RESULT 14

AAAB12570 standard; Protein: 449 AA.

AAAB12570;

09-NOV-2000 (first entry)

NDO related complex alpha subunit protein sequence SEQ ID NO:18.

Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO; inducible multi-component enzyme; alpha subunit; mutant; chiral diol; polymer; resin; pharmaceutical; rubber industry; bioremediation.

Pseudomonas sp. Synthetic.

WO200037480-A1.

29-JUN-2000.

26-OCT-1999; 99WO-US25079.

26-OCT-1998; 98US-0105575.

(IOWA ) UNIV IOWA RES FOUND.

Parales R, Gibson D, Resnick S, Lee K;

WPI: 2000-452174/39.

N-PSDB: AAA65344.

Novel naphthalene dioxygenase mutant having a specific amino acid substitution for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry and for carrying out bioremediation



```
QY 121 KDIYGSINKKCIJGKRVAVESFHFYIGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL 180
|:|||||
Db 121 KELYGESLNKCKLGKLEVARVESFHFYIGCFODEAPSLMDYLGDAAMYLEPIFKHSGGL 180
|:|||||
QY 181 ELVGPBGKVVYIKANMKAPAEFNVGDAYHVGWTHASSLSRSGESIFSSLAGNALPPEGAGL 240
|:|||||
Db 181 ELVGPBGKVVYIKANMKAPAEFNVGDAYHVGWTHASSLSRSGESIFSSLAGNALPPEGAGL 240
|:|||||
QY 241 QMTSKYSGMGVLMGYSVHSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNCV 300
|:|||||
Db 241 QMTSKYSGMGVLMGYSVHSADLVPELMAFGAKOERLNKEIGDVARIYRSHLNCV 300
|:|||||
QY 301 FPNNSMLTCSGVFPKWNPIDANTTEVWTYATVEKMDPEDLKRLADSVORTLGPAFWES 360
|:|||||
Db 301 FPNNSVLTCSGVFPKWNPIDANTTEVWTYATVEKMDPEDLKRLADAVORTVGPAGFWES 360
|:|||||
QY 361 DDNDNMETASQNGKKYOSRSDLSNLGFGEDVYGDVAYPGVGSALGETSYRGFYRAY 420
|:|||||
Db 361 DDNDNMETASQNGKKYOSRSDLSNLGFGKDYGDVAYPGVGSALGETSYRGFYRAY 420
|:|||||
QY 421 QAHVSSSNMAFEHASSTWHTELTKTDR 449
|:|||||
Db 421 QAHVSSSNMAFEHASSTWHTELTKTDR 449
|:|||||
```

Search completed: January 27, 2003, 08:58:51  
Job time : 30.1667 secs



```

Matches 165: Conservative 64: Mismatches 174: Indels 50: Gaps 8
Oy 20 IGHDEELPQHEKLTIFARWMLFLTHDSLIPAGDVTYAMGIDEVIVSQNQNGSIRAFIN 79
Db 24 IYDPIIDYIELERFEGRCWFLAHESQILPKGDFENITMGEDAVVYVROKDGSIKAFIN 83
Oy 80 VCRHRGKTLVSYEAGNAKQFVCSYHGWMGFGSNGELOSYPFEKDLYGESLNKKCLGEYA 139
Db 84 QCRHRAMRVSYADCNGTRAFCTCPHYQMSYGINIGELIDVLEPRAPYQGLCKSHWGLNEVP 143
Oy 140 RVESHGFLTGCFODEAPPLIMDYLDGDAVLEPMF-KHSGGLELVGPCKVYIKAMKAP 198
Db 144 CVESYKGLFEGMWDYAPDLRDYLDGDIAYLDGMIDRRGGTGEIVGGYQKWVINCWKPEP 203
Oy 199 AENFYGDAVHYWMTBASL-----RSGESIFSSLGNALP----- 234
Db 204 AEGFASDQYHALFASHANQVVLGAKDDGSDKRLGD-----GQPARPWEHTAKDALQFG 256
Oy 235 --PEGAGLQMTSKYSGMGVLMDYGSVGHASDLVPELMAFGAKOERLNKEIGDVARIY 292
Db 257 QDHGSGSEFFETKPPANVWV--DGAVSSYRYRETYAE-----AEQRAGEVYALTL 303
Oy 293 RSHLQCTVPEPNSMLTSCGVFKWNPIDANTNEVMTYALVERKMPEDLKRRLADSVQRTI 352
Db 304 AGHNH--IFPTLSWLNGLTITLAVMHRGRGDQYEVNAFCLTIDAASDVAAAEENSITRAF 361
Oy 363 GPAGWESDNDNMETASONGKKYOSRSDLSLNLGFGEDVYGDVAYVPGVAKSAIGETS 412
Db 362 GPAGLEDDDSSENCEIOLKLGHRNRNSKLCLEMGILGQEKRRDDIGPIT-NYIFSETA 420
Oy 413 YRGFYRAYOAHVSSNNMAEFEEHASSPWHTELTK 445
Db 421 ARGMYQRNADLLSSSEMQVELDKTAAYQGEVAK 453

RESULT 2
US-09-738-626-6140
: Sequence 6140, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAOKO
: APPLICANT: SENOH, AKIHIRO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738, 626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 6140
: LENGTH: 490
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
: US-09-738-626-6140

```

Query Match	15.7%	Score 377;	DB 9;	Length 490;
Best Local Similarity	26.3%;	Pred. No. 1.3e-25;		
Matches 118;	Conservative 66;	Mismatches 215;	Indels 50;	Gaps 15;
0Y	20	IHGDEELFQHEHLEKTFIARNWLEFLTHDSLLIPARGDVTYAKMGIDEIVYSRQNDGSTRAFLN		
		79		

[illegible]

```

RESULT 3
US-09-815-242-11692
Sequence 11692, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELTRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11692
LENGTH: 385
TYPE: PRF
ORGANISM: Klebsiella pneumoniae
US-09-815-242-11692

```

Query Match	15.68; Score 376.5; DB 10; Length 385;
-------------	--



Best Local Similarity 31.0%; Pred. NO. 1e-25;  
Matches 110; conservative 58; Mismatches 154; Indels 33; Gaps 16

Oy	34	IFARMLITLDSLPDPAVDVTKMGIDEVYERODGSTRATLNCRRKRGVSYEA	93
Db	3	IFEGMFWLAEHSIIPQPGDYITLLTGROPYIITRDKKNELHALINSCAIRGAMLCRKRT	62
Oy	94	GNAKGFVCSYHGMGPGNGELQSVPEEK-DLYGESL-NKKCLGLEKVAVESFHEGYGC	151
Db	63	GNKNSKTFPEFGMFWFSNNKGLKLAKKDESTGAYRPTFFKHESHDLOKPRQSYKGLFSGS	122
Oy	152	FDOEARPLMDVLGDAAWLEBPMFKHS-GGELVPRPKVYIKAWKPAEENFYODAVHYG	210
Db	123	LNADQPLFENLTGCTCKIDILYDQAREGDEVLKGGSSYYEGMKWICGAEN-GADGYHVS	184
Oy	211	---WTHASSL--RGESITFSLSAGNAALDPBGAGLQMTSKYG--SGMGVLMGDSGVHSA	263
Db	182	VVHNMYASTMSRRNYEA----EGTHVTVDANGMSKSLGGGGEFGNGHMLLW-----TR	229
Oy	264	DLVPELMAFGAKOERLKEIGDVRA-RYRSHLNCIVFPNMSML-THCSGFKVKKWIDA	321
Db	230	ALNPEVRVY-YAHERPLQAEGERRADQMMENFNNLLCYLNVYIMDFSIQIYIRIAV	288
Oy	322	NTTEY--WTYAIIV-EKMPEDLKRRLADSVORTGPRGMESSDDNDNMETASONG	373
Db	289	DKTEYITLWCEAPKESDQARLRIIR--QYEDFFNYSGMGCPDLEEF-SACORG	339

```

RESULT 4
US-09-815-242-5097
Sequence 5097, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: zykind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
CURRENT FILING DATE: 2001-03-21
PRIORITY APPLICATION NUMBER: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY APPLICATION NUMBER: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY APPLICATION NUMBER: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY APPLICATION NUMBER: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY APPLICATION NUMBER: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY APPLICATION NUMBER: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5097
LENGTH: 424
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5097

```

Query Match	14.2%	Score 341.5;	DB 10;	Length 424;
Best Local Similarity	27.6%;	Pred. No. 1.5e-22;		
Matches 115;	Conservative 70;	Mismatches 177;	Indels 55;	Gaps 16;

[illegible]

RESULT 5  
 US-09-776-490-19  
 Sequence 19, Application US/09776490  
 Patent No. US20010012886a1  
 GENERAL INFORMATION:  
 APPLICANT: Brilgs, Steven P.  
 Jihal, Gurmukh S.  
 Gray, John  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
 CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
 NUMBER OF SEQUENCES: 65  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
 STREET: P.O. Drawer 34009  
 CITY: Charlotte  
 STATE: No. US20010012886a1th Carolina  
 COUNTRY: USA  
 ZIP: 28234  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/776,490  
 FILING DATE: 02-Feb-2001  
 CLASSIFICATION: <Unknown>  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/810,009  
 FILING DATE: 04-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Spruill, W. Murray  
 REGISTRATION NUMBER: 32,943  
 REFERENCE/DOCKET NUMBER: 5718-4  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-881-3140  
 TELEFAX: 919-881-3175  
 TELEX: 575102  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 35 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: <unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-490-19

Query Match 7.7%: Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%; Pred. No. 3.1e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCHRGKTLVSEAGNAKGFVCSYHGWFSGSNGE 113  
|||||  
Db 1 NVCHRGKTLVSEAGNAKGFVCSYHGWFSGSNGK 35

RESULT 6  
US-09-776-491-19  
Sequence 19, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1ch Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776.491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810.009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32.943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-776-491-19

Query Match 7.7%: Score 185; DB 10; Length 35;  
Best Local Similarity 94.3%; Pred. No. 3.1e-10;  
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCHRGKTLVSEAGNAKGFVCSYHGWFSGSNGE 113  
|||||  
Db 1 NVCHRGKTLVSEAGNAKGFVCSYHGWFSGSNGK 35

RESULT 7  
US-09-776-490-20  
Sequence 20, Application US/09776490  
Patent No. US20010012886A1

GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010012886A1ch Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776.490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810.009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32.943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-490-20

Query Match 7.2%: Score 174; DB 10; Length 35;  
Best Local Similarity 85.7%; Pred. No. 3e-09;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCHRGKTLVSEAGNAKGFVCSYHGWFSGSNGE 113  
|||||  
Db 1 NVCHRGKTLVSEAGNAKGFVCSYHGWFSGSNGK 35

RESULT 8  
US-09-776-491-20  
Sequence 20, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1ch Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 20:  
US-09-776-491-20

Query Match  
Best Local Similarity 7.2%; Score 174; DB 10; Length 35;  
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DB 1 NVCRHRGKTLVDAEAGNAGKPGVCGYHGWGSGNSGK 35

RESULT 9  
US-09-776-490-21  
Sequence 21, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140

TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-490-21

Query Match  
Best Local Similarity 7.0%; Score 168; DB 10; Length 35;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DB 1 NVCRHRGKTLVDAEAGNAGKPGVCGYHGWGSGNSGK 35

RESULT 10  
US-09-776-491-21  
Sequence 21, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: NO. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 21:  
US-09-776-491-21

Query Match  
Best Local Similarity 7.0%; Score 168; DB 10; Length 35;  
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NCRHRGKTLVSEAGNAKGFVCSYHGNGFSGSNE 113  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 NCRHRGKTLVDAGNAKGFVCSYHGNGFSGSNGK 35

## RESULT 11

US-08-976-063C-4  
; Sequence 4, Application US/08976063C  
; Publication No. US20020182697A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Steinbuechel; Horst Priefert; Jurgen Rabenhorst  
; TITLE OF INVENTION: SYNTHETIC ENZYMES FOR THE PRODUCTION OF  
; TITLE OF INVENTION: CONFERTIL ALCOHOL, CONFERYLALDEHYDE, FERULIC ACID, VANILLIN A  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPRUNG KRAMER SCHAEFFER & BRISCOE  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-5144  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB storage  
; COMPUTER: HP VECTOR  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/976,063C  
; FILING DATE: 21-NOV-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 196 49 655.1 (Germany)  
; FILING DATE: 29-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141  
; REFERENCE/DOCKET NUMBER: Bayer 9998-CAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 332-1700  
; TELEFAX: (914) 332-1844  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 354 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-976-063C-4

Query Match 4.9%; Score 118.5; DB 1; Length 354;  
Best Local Similarity 23.2%; Pred. No. 0.0068;  
Matches 56; Conservative 40; Mismatches 78; Indels 67; Gaps 14;

QY 34 IFARN-WLFLTHDSLIPAGQDIYTAAMGIDVYVRONDSITAFNLVCHRGKTLVSV 91  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 MFKMAMVYACTPDELIDKP---LGRQICNEKIVFYKGPBGRVAVEDFCFHRGAPL--- 54  
QY 92 EAGNAKGFV-----CSYHGNGFSGSNGELOSVPFEKDLVYSESLKKKCG--LKEVARVE 142  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 55 -----SLGFVADGKILICYHGLEMGCEGKTLAMP-----GQRVOGFPCITASYA-VE 99  
QY 143 SFHGFTY--GCFDQEARPLMDYLGDAAWYLEPMFKHSGGLELVGPGKVVIRKANWKAPA 199  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 100 ERYGFTIWMVGDRELADRALIHL---EMADNPEMAYGGGL-----YHADIYRLMI 148  
QY 200 ENFVGDAIVGWMTHASLR-----SGESIFSS-LAGNALP-----EGAG 239  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 149 DNLM-DLTHETVYHASSIGKEIDEAPVSTRVEDTVITSRMDNVAMAPPEWRAALRGNG 207  
QY 240 L 240  
Db 208 L 208

## RESULT 12

US-09-776-490-12  
; Sequence 12, Application US/09776490  
; Patent No. US2001001286A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; John, Gurmukh S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009  
; CITY: Charlotte  
; STATE: No. US2001001286A1th Carolina  
; COUNTRY: USA  
; ZIP: 28234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/776,490  
; FILING DATE: 02-Feb-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/810,009  
; FILING DATE: 04-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spruill, W. Murray  
; REGISTRATION NUMBER: 32,943  
; REFERENCE/DOCKET NUMBER: 5718-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-881-3140  
; TELEFAX: 919-881-3175  
; TELEX: 575102  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-490-12

Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0014;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 79 NCRHRGKTLVSEAGNAKGFVCSYHGNGFSGSNG 112  
| | | | | : | | | | | : | | | | | : | | | | | :  
Db 1 NCRHRGMRICRADAGNAKAPGCSYHGMAVDYTAG 34  
RESULT 13  
US-09-776-491-12  
; Sequence 12, Application US/09776491  
; Patent No. US20010013135A1  
; GENERAL INFORMATION:  
; APPLICANT: Briggs, Steven P.  
; John, Gurmukh S.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
; CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
; STREET: P.O. Drawer 34009

CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-776-491-12  
Query Match 4.6%; Score 110; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0014;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;  
QY 79 NVCRHGRKTLVSEAGNAKGFVCSYHGWFSGSNG 112  
1 | | | | | : | | | | | | | | | | : | |  
Db 1 NQCRHGRKICRADGNNAKAPTCSTHGWAYDSAG 34  
RESULT 14  
US-09-776-490-14  
Sequence 14, Application US/09776490  
Patent No. US20010012886A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010012886A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,490  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
US-09-776-490-14  
Query Match 4.5%; Score 109; DB 10; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0017;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;  
QY 79 NVCRHGRKTLVSEAGNAKGFVCSYHGWFSGSNG 112  
1 | | | | | : | | | | | | | | | | : | |  
Db 1 NQCRHGRKICRADGNNAKAPTCSTHGWAYDSAG 34  
RESULT 15  
US-09-776-491-14  
Sequence 14, Application US/09776491  
Patent No. US20010013135A1  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
Johal, Gurmukh S.  
Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. US20010013135A1th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/776,491  
FILING DATE: 02-Feb-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/810,009  
FILING DATE: 04-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-776-491-14

Query Match 4.5%; Score 109; DB 10; Length 35;  
 Best Local Similarity 52.9%; Pred. No. 0.0017;  
 Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY 79 NVCRHRCITVSEAGNAKGFVCSYHGWFSGSNG 112  
 | | | | | : | | | | | | | | | : | |  
 DB 1 NCCRRHGRICRADGNGNAKSPTCSYHGWAYDSAG 34

Search completed: January 27, 2003, 09:20:58  
 Job time : 8 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:56:07 : Search time 10.1667 Seconds  
(without alignments)  
1299.432 Million cell updates/sec

Title: US-09-843-250-36

Perfect score: 2408

Sequence: 1 MNYNKILVESGLSKHLI.....AEFEHASTWHTLTXTTDR 449

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2-6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2-6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2-6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2-6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2-6/ptodata/1/1aa/PTCUTS\_COMB.pep:\*  
6: /cgn2-6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	224	9.3	446	US-09-004-393B-4	Sequence 4, Appl1
2	223	9.3	439	US-09-004-393B-2	Sequence 2, Appl1
3	185	7.7	35	US-08-810-009-19	Sequence 19, Appl1
4	174	7.2	35	US-08-810-009-20	Sequence 20, Appl1
5	168	7.0	35	US-08-810-009-21	Sequence 21, Appl1
6	112.5	4.7	379	US-09-028-934-36	Sequence 36, Appl1
7	110	4.6	35	US-08-810-009-12	Sequence 12, Appl1
8	109	4.5	35	US-08-810-009-14	Sequence 14, Appl1
9	108	4.5	35	US-08-810-009-13	Sequence 13, Appl1
10	108	4.5	35	US-08-810-009-15	Sequence 15, Appl1
11	108	4.5	622	US-09-311-626B-4	Sequence 4, Appl1
12	106	4.4	35	US-08-810-009-9	Sequence 9, Appl1
13	104	4.3	35	US-08-810-009-18	Sequence 18, Appl1
14	103.5	4.3	432	US-08-809-326A-16	Sequence 16, Appl1
15	103.5	4.3	649	US-08-809-326A-15	Sequence 15, Appl1
16	103	4.3	35	US-08-810-009-11	Sequence 11, Appl1
17	102	4.2	35	US-08-810-009-8	Sequence 8, Appl1
18	102	4.2	35	US-08-810-009-10	Sequence 10, Appl1
19	101	4.2	17	US-08-810-009-44	Sequence 44, Appl1
20	101	4.2	17	US-08-810-009-45	Sequence 45, Appl1
21	100	4.1	35	US-08-810-009-17	Sequence 17, Appl1
22	99.5	4.1	256	US-09-325-932A-57	Sequence 57, Appl1
23	95	3.9	17	US-08-810-009-46	Sequence 46, Appl1
24	94.5	3.9	350	US-08-828-922-1	Sequence 1, Appl1
25	94	3.9	560	US-08-814-052-6	Sequence 6, Appl1
26	94	3.9	560	US-08-812-829-6	Sequence 6, Appl1
27	93	3.9	35	US-08-810-009-16	Sequence 16, Appl1

28	93	3.9	421	US-09-239-303-2	Sequence 2, Appl1
29	93	3.9	563	US-09-134-001C-4800	Sequence 4800, Ap
30	93	3.9	5588	US-09-036-987A-6	Sequence 6, Appl1
31	93	3.9	5588	US-09-370-700-6	Sequence 6, Appl1
32	92	3.8	1087	US-08-570-311-8	Sequence 8, Appl1
33	92	3.8	1087	US-08-353-485-8	Sequence 8, Appl1
34	92	3.8	1358	US-08-570-311-27	Sequence 27, Appl1
35	91.5	3.8	446	US-07-627-539G-7	Sequence 7, Appl1
36	91.5	3.8	468	US-07-627-539G-2	Sequence 2, Appl1
37	91	3.8	1693	US-08-478-507-7	Sequence 7, Appl1
38	91	3.8	1693	US-09-128-275A-7	Sequence 7, Appl1
39	91	3.8	1693	US-09-553-427-7	Sequence 7, Appl1
40	90.5	3.8	345	US-09-362-133A-2	Sequence 2, Appl1
41	90.5	3.8	346	US-09-149-476-493	Sequence 493, App
42	90.5	3.8	525	US-08-348-891A-2	Sequence 2, Appl1
43	90.5	3.8	525	US-08-905-817-2	Sequence 2, Appl1
44	90.5	3.8	774	US-09-276-400-7	Sequence 7, Appl1
45	90.5	3.8	774	US-09-448-076-7	Sequence 7, Appl1

#### ALIGNMENTS

```
RESULT 1
US-09-004-393B-4
: Sequence 4, Application US/09004393B
: Patent No. 6310271
: GENERAL INFORMATION:
: APPLICANT: Hanson D., Andrew
: APPLICANT: Rathinasabapathi, Bala
: TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
: FILE REFERENCE: US-162
: CURRENT FILING DATE: 1998-01-08
: PRIOR APPLICATION NUMBER: 60/035,147
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 4
: LENGTH: 446
: TYPE: PRT
: ORGANISM: Beta vulgaris
US-09-004-393B-4

Query Match          9.3%, Score 224; DB 4; Length 446;
Best Local Similarity 27.4%; Pred. No. 6e-13;
Matches 58; Conservative 37; Mismatches 97; Indels 20; Gaps 6;

QY 11 ESGLSOKHLIHODELFOHELKTIFARWMLFLTHSLIPAPEDYVYATKMGIDIVYSRON 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 99 EDALPPSTWYTEPFYSHLEIRIFYKGOVAGYSEOVKEKNQYTGSGNVEYLVSBDG 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 71 DGSIAFNVCHNRKTLVSVAGNAKGFVCSYHMGFGSGNGELQSVPEFKLVGESLNK 130
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 159 OGELAFNVCTHRA-SILACSGKKSCTVCPYHGWYVGLDGLSLAKA--SKATETQNDLP 215
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 131 KCLGKEVARVESFNGFYGCEDQAPPLMD---YLDA-----AWYLEPMFKHSGGLE 181
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 216 KELGIAPL-KVAEMQPFILISLDRSLDANADVGTWICKSADVYKAHAFDPLKATTHRSE 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 LVGPPGKVVIRANKAPAEENFYGDAYHYGWTH 213
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 F-----PMECMKVFCFDNYLDSSYHVPYAH 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
US-09-004-393B-2
: Sequence 2, Application US/09004393B
: Patent No. 6310271
: GENERAL INFORMATION:
: APPLICANT: Hanson D., Andrew
```

```
; APPLICANT: Rathnasabapathi, Bala
; APPLICANT: Burnet, Michael
; TITLE OF INVENTION: Polynucleotides Encoding Choline Monooxygenase and
; TITLE OF INVENTION: Plants Transformed Therewith
; FILE REFERENCE: UF-162
; CURRENT APPLICATION NUMBER: US/09/004,393B
; PRIOR FILING DATE: 1998-01-08
; PRIOR APPLICATION NUMBER: 60/035,147
; PRIOR FILING DATE: 1997-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Spinacia oleracea
US-09-004-393B-2

Query Match
Best Local Similarity 9.3%; Score 223; DB 4; Length 439;
Matches 58; Conservative 38; Mismatches 83; Indels 14; Gaps 7;

QY 27 FQHELTIFARNMLFLHDSLPAPGDYVTAKKGIDVIVSRQDGSRAFTLNCRRGK 86
DB 108 YSHELRTIFKQVAGISQIKEPNOYFTGSGNVEYLVSRDGEKVAFHNVCTHRA- 166
QY 87 TLVSVENAGNAKGFVCSYHGFGSNGELQSVPEKDLGSLNKKCLGKEVARVESFHG 146
DB 167 SILAGSGKSKSCVCPHYGVGMDSGLAKASKAKP--EQNLDPKELGIYPL-KVAVWGP 223
QY 147 FYICCDQEAAPLMD---YLGDAAWYLEPMFKHS--GGLIELVGPCKVKVIAKWKAPAE 200
DB 224 FVLISDRSLBEGCDVGTETWLGTSA--EDVKAFAPDPSLQFI-HRSEFPWESNMKIFSD 279
QY 201 NEVGDAYHVGWTH 213
DB 280 NYDSSIHVPYAH 292

RESULT 3
US-08-810-009-19
; Sequence 19, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
```

```
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-810-009-19

Query Match
Best Local Similarity 7.7%; Score 185; DB 4; Length 35;
Matches 33; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 79 NVCHRGRKTLVSVENAGNAKGFVCSYHGFGSNGE 113
DB 1 NVCHRGRKTLVSVENAGNAKGFVCSYHGFGSNGK 35

RESULT 4
US-08-810-009-20
; Sequence 20, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-810-009-20

Query Match
Best Local Similarity 7.2%; Score 174; DB 4; Length 35;
Matches 30; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 79 NVCHRGRKTLVSVENAGNAKGFVCSYHGFGSNGE 113
DB 1 NVCHRGRKTLVSVENAGNAKGFVCSYHGFGSNGK 35

RESULT 5
US-08-810-009-21
```



```
; Sequence 21, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Jomal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P. O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprull, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-21

Query Match          7.0%; Score 168; DB 4; Length 35;
Best Local Similarity 80.0%; Pred. No. 2.6e-09;
Matches 28; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 79 NVCRRHGRKTVSEAGNAGKGFVCSYHMGFGSGNGE 113
Db 1 NVCRRHGRKTVSEAGNAGKGFVCSYHMGFGSGNGK 35

RESULT 6
US-09-028-934-36
; Sequence 36, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Iigon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyroinhibitin Biosynthesis Genes and Uses
; TITLE OF INVENTION: Theoreof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028.934
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC1506/CIP7
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-028-934-36

Query Match          4.7%; Score 112.5; DB 3; Length 379;
Best Local Similarity 20.3%; Pred. No. 0.019;
Matches 95; Conservative 54; Mismatches 149; Indels 171; Gaps 26;

QY 20 IHGDEELFQ----HELKTIFARNW-LFLTHSLIPAPGDVYTAKKGIDEIVTSRONDS 73
Db 5 IHQEPRIQASGVNDLTQTQTAASWYVAMRSDALRKP---VAIKLFGQPLVAMRDGGGR 61

QY 74 IRAFLNVCRRHGRKTVSEAGNAGKGFV-----CSYHMGFGSGNGELQSVPEKDLXGS 127
Db 62 PYVMERYCSHLGASH-----AKGVVEGCIQCPHNMRYDSTGACSHVP-----GHS 108

QY 128 LKKKICGLKEVARVESF-----HGFTGCFDQEAR---PLMDYLGDAMY 169
Db 109 -----TEVRLPEIIPPTARQSVYPMERYGFVWWTGTRAPLPLEMPAESESEH 160

QY 170 LPEMFHSGGLLEVPBGKVVITKANWKAPAEFVGDAYHYGWTBASSLSGE-SIF---S 225
Db 161 QSLRFAYETTTSVL-----RII-----ENFY-DAQHAAPVHQLPISAFELKLPDESS 206

QY 226 SLAGNALPPE---GAGIOM-TSKYSGMGVL-----WDGYSGVHSAADLV 266
Db 207 PPPGQALARDGAMFGAGIDFHVDRYFGPLGVISRTLGISMSRMQLHFGDYPG----- 259

QY 267 PELMAFGAKOERLNKEIDVRAIRYRSHLNCYVFPNNSMLFCSGVFKVWNPIDANTTEV 326
Db 260 -----GCTMYSLD---GDVKTRL---LQCV-----TPYDKETIWA 289

QY 327 WTYAIVKMDPBDLKRRLADSV---ORTTGPAGF---WESDNDNMETASONGKKYQSR 379
Db 290 HHLLAIKKG--DGVRSANFLIYGLQTTWAAGDYVAIINSMKAD-----GGGAFSKY 340

QY 380 DSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFTYAYQAHYSSN 428
Db 341 DDLIL-----KYRAFRRWVKVALEN 362

RESULT 7
US-08-810-009-12
; Sequence 12, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
```

APPLICANT: Johal, Gurmukh S.  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-12

Query Match 4.6%; Score 110; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.00086;  
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

OY 79 NVCRRGKTLVSVEAGNAKGVCSTYHGWFSGSNG 112  
| | | | | : : | | | | | | | | : | |  
Db 1 NCCRHRGMRICRADAGNAKAPTCSYHGWAYDTAG 34

RESULT 8  
US-08-810-009-14  
Sequence 14, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-12

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-14

Query Match 4.5%; Score 109; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0011;  
Matches 18; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

OY 79 NVCRRGKTLVSVEAGNAKGVCSTYHGWFSGSNG 112  
| | | | | : : | | | | | | | | : | |  
Db 1 NCCRHRGMRICRADAGNAKAPTCSYHGWAYDSAG 34

RESULT 9  
US-08-810-009-13  
Sequence 13, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johal, Gurmukh S.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-13

Query Match 4.5%; Score 108; DB 4; Length 35;  
Best Local Similarity 52.9%; Pred. No. 0.0013;

```
Matches 18; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
QY 79 NVCRRHGKTLVSEAGNAGKGFVCSYHGWFSGSNG 112
| | | | | : | | | | | | | | | : |
Db 1 NCCRRHGMRCISDAGNNAKAPFCSTYHGWAIDYDNG 34

RESULT 10
US-08-810-009-15
; Sequence 15, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 5718-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-881-3140
; TELEFAX: 919-881-3175
; TELEX: 575102
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-810-009-15

Query Match 4.5%; Score 108; DB 4; Length 35;
Best Local Similarity 50.0%; Pred. No. 0.0013;
Matches 17; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
QY 79 NVCRRHGKTLVSEAGNAGKGFVCSYHGWFSGSNG 112
| | | | | : | | | | | | | | | : |
Db 1 NCCRRHGMRCISDAGNNAKAPFCSTYHGWAIDYDNG 34

RESULT 11
US-09-311-626B-4
; Sequence 4, Application US/09311626B
; Patent No. 6399347
; GENERAL INFORMATION:
; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schroff, Kirk
; APPLICANT: Andersen, Lene No. 6399347boe
; APPLICANT: Schuelein, Martin
; APPLICANT: Outtrup, Helle
; TITLE OF INVENTION: No. 6399347el Rhamnogalacturonan Hydrolases
; FILE REFERENCE: 5572.204-US
; CURRENT APPLICATION NUMBER: US/09/311.626B
```

```
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: 0608/98
; PRIOR FILING DATE: 1998-05-01
; PRIOR APPLICATION NUMBER: 60/084,358
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: prt
; LENGTH: 622
; ORGANISM: Bacillus licheniformis
US-09-311-626B-4

Query Match 4.5%; Score 108; DB 4; Length 622;
Best Local Similarity 21.6%; Pred. No. 0.11;
Matches 90; Conservative 52; Mismatches 137; Indels 138; Gaps 23;
QY 84 RCKTLVSEAG--NAKGFVCS-----YHGWFSGSNGELQSVPEK-----DLYG 125
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 245 KGVIGNNANADYRNAQGRILSGPEYLVFKG--DTGAELTVNVEPARGNVADMGDSYG 301
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 126 ESLNKKCLGL-----KEVARVESFHGFTIG-----CFDQEARPLMD 161
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 302 NKVDREFLAGVAYLDGERSEFVWARGYTRTVLVAYNFRG--GKLTKLMTFSDAPGNGA 358
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 162 YLGDAWYLEPMFKHSGLELVGPPGVYIKANWKAPAEFVDAYHYGWTFASSLSRGE 221
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 359 YAGOG-----NHSLVADYDGDGKDEIITYGAAYVDHDKG-LYSTGWHGDMHNGN 409
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 222 STFSSLAGNALPPEGAGLQMTSKYSGMGVYLDGYSGVHSAADVPELMAFGAKQERL- 280
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 410 -----LDPSRPGLEVFO-----VHENSNSPYGLSPFADAKTKI 443
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 281 -----NKEIG-DYRARIRYSHLNCVFPNNMSLTCGYEKKVNPIDANTT-EVWTYALVE 333
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 444 WGVHAKGDVGRGMADIDPRYEGAEVWANGSLYTAQV-KIGNTLPSSNFGIW----- 496
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 334 KDMPEDLKRLADSVQRTIGPAGFWESDND--NMETAS---ONG-KKYOSRSDLSLN 386
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 497 --WGDLDRELDS-----NRIDKMDYONSRTVNLITLSGASANNGTATPDSLADILCD 549
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 387 -----LGFGEED-----VYDAVYPGVYKSAIG--ETSYRGVY 417
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 550 WREEVWRAEDSELSRLYTTTDDVTEHRMYTLMHDAVYRLGIAMQVGNQPHRTGFY 606
| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 12
US-08-810-009-9
; Sequence 9, Application US/08810009
; Patent No. 6211437
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Johal, Gurmukh S.
; APPLICANT: Gray, John
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING
; TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BELL, SELTZER, PARK & GIBSON
; STREET: P.O. Drawer 34009
; CITY: Charlotte
; STATE: No. 6211437th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810.009
; FILING DATE: 04-MAR-1997
; CLASSIFICATION: 800
```

ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-9

Query Match 4.48; Score 106; DB 4; Length 35;  
Best Local Similarity 51.48; Pred. No. 0.0021;  
Matches 18; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 79 NVCRHRCGLTVSEAGNAKGFVCSYHGMGSGSNGE 113  
DB 1 NSCRHRCGLTVSEAGNAKGFVCSYHGMGSGSNGE 35

RESULT 13  
US-08-810-009-18  
Sequence 18, Application US/08810009  
Patent No. 6211437  
GENERAL INFORMATION:  
APPLICANT: Briggs, Steven P.  
APPLICANT: Johnal, Gurmukh S.  
APPLICANT: Gray, John  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR CONTROLLING  
TITLE OF INVENTION: CELL DEATH AND DISEASE RESISTANCE IN PLANTS  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BELL, SELTZER, PARK & GIBSON  
STREET: P.O. Drawer 34009  
CITY: Charlotte  
STATE: No. 6211437ch Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,009  
FILING DATE: 04-MAR-1997  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 5718-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-881-3140  
TELEFAX: 919-881-3175  
TELEX: 575102  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-810-009-18

Query Match 4.38; Score 104; DB 4; Length 35;  
Best Local Similarity 52.98; Pred. No. 0.0032;  
Matches 18; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

QY 79 NVCRHRCGLTVSEAGNAKGFVCSYHGMGSGSNGE 112  
DB 1 NSCRHRCGLTVSEAGNAKGFVCSYHGMGSGSNGE 34

RESULT 14  
US-08-809-326A-16  
Sequence 16, Application US/08809326A  
Patent No. 6165478  
GENERAL INFORMATION:  
APPLICANT: Izutsu, Hiroshi  
APPLICANT: Obara, Kazuhiko  
APPLICANT: Matsumoto, Akira  
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,  
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING  
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA  
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O  
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT  
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA  
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND  
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND  
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,326A  
FILING DATE: 19-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 224711/94  
FILING DATE: 20-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106006/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106008/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106009/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106010/95  
FILING DATE: 28-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 106011/95  
FILING DATE: 28-APR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Miller, Charles E.  
REGISTRATION NUMBER: 24,576  
REFERENCE/DOCKET NUMBER: 7426-043-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:54:32 ; Search time 25.8333 Seconds  
(without alignments)  
3581.232 Million cell updates/sec

Title: US-09-843-250-36  
Perfect score: 2408  
Sequence: 1 MNYNKKILVSESGISQKHLI.....AEPFHSSFTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_potent:\*  
13: sp\_virus:\*  
14: sp\_unclassified:\*  
15: sp\_rvirs:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2332	96.8	449	2	Q938R9
2	2292	95.2	449	2	Q52142
3	2208	91.7	449	2	Q92173
4	2193	91.1	447	2	Q915D2
5	2189	90.9	447	2	Q52382
6	2074	86.1	447	2	Q8VUD4
7	2056	85.4	447	2	P95564
8	2027	84.2	447	2	Q8RTT4
9	1967.5	81.7	451	2	Q45695
10	1490	61.9	277	2	Q9ETK2
11	1488	61.8	277	2	Q9F554
12	1486	61.7	277	2	Q9F573
13	1484	61.6	277	2	Q9F558
14	1483	61.6	277	2	Q9F559
15	1451	60.3	277	2	Q9F553
16	1405	58.3	277	2	Q9EUC7

17	1403	58.3	277	2	Q9F578	Q9F578	pseudomonas
18	1402	58.2	277	2	Q9F575	Q9F575	pseudomonas
19	1399	58.1	277	2	Q9ETV3	Q9ETV3	pseudomonas
20	1399	58.1	277	2	Q9F555	Q9F555	pseudomonas
21	1396	58.0	277	2	Q9F576	Q9F576	pseudomonas
22	1396	58.0	277	2	Q9F557	Q9F557	pseudomonas
23	1393	57.8	277	2	Q9F574	Q9F574	pseudomonas
24	1393	57.8	277	2	Q9F572	Q9F572	pseudomonas
25	1393	57.8	277	2	Q9F571	Q9F571	pseudomonas
26	1393	57.8	277	2	Q9F570	Q9F570	pseudomonas
27	1393	57.8	277	2	Q9F566	Q9F566	pseudomonas
28	1390	57.7	277	2	Q9F577	Q9F577	pseudomonas
29	1384	57.5	277	2	Q9F552	Q9F552	pseudomonas
30	1355.5	56.3	450	2	Q9ZHH3	Q9ZHH3	burkholderi
31	1139	47.3	226	2	Q8VUM6	Q8VUM6	raistonia s
32	1134	47.1	214	2	Q8VR23	Q8VR23	pseudomonas
33	1134	47.1	214	2	Q8VL21	Q8VL21	pseudomonas
34	1127	46.8	214	2	Q8VR24	Q8VR24	pseudomonas
35	1123	46.6	214	2	Q8VR22	Q8VR22	pseudomonas
36	1113.5	46.2	455	2	Q93NA8	Q93NA8	burkholderi
37	1067	44.3	214	2	Q8VR25	Q8VR25	burkholderi
38	1045.5	43.4	437	2	Q9WXG8	Q9WXG8	alcaligenes
39	1040	43.2	208	2	Q9F6B6	Q9F6B6	marinobacte
40	1039	43.1	206	2	Q8VUM5	Q8VUM5	naphthalene
41	1038	43.1	214	2	Q8VR21	Q8VR21	burkholderi
42	1038	43.1	303	2	Q8VUM2	Q8VUM2	raistonia s
43	1016	42.2	297	2	Q93M40	Q93M40	uncultured
44	1009	41.9	297	2	Q93M39	Q93M39	uncultured
45	1002	41.6	297	2	Q93M41	Q93M41	uncultured

#### ALIGNMENTS

RESULT 1	ID	Q938R9	PRELIMINARY;	PRT;	449 AA.
AC	Q938R9				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)			
DE	Naphthalene dioxygenase.				
GN	NAHAC.				
OS	Pseudomonas fluorescens.				
OC	Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;				
CC	Pseudomonas.				
OX	NCBI TaxID=294;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Min K.-H., Ji S.-H.;				
RL	Submitted (CGT-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AY048759; AL07262.1; -.				
DR	InterPro: IPR001281; Rleske.				
DR	InterPro: IPR001663; Ring_hydroxyl_A.				
DR	Pfam: PF00355; Rleske; 1.				
DR	Pfam: PF00848; Ring_hydroxyl_A; 1.				
DR	PROSITE: PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.				
KW	Dioxygenase.				
SQ	SEQUENCE 449 AA: 49608 MW: 55FEDBE0282FE812 CRC64;				
Query Match	Best Local Similarity	96.8%; Score 2332; DB 2; Length 449;			
Matches	431; Conservative	12; Mismatches	6; Indels	0; Gaps	0;
QY	1	MNYNKKILVSESGISQKHLIHGDEELFOHELKTFARNWLFTHDSLIAPAGDYVTAKMG	60		
DB	1	MNYNKKILVSESGISQKHLIHGDEELFOHELKTFARNWLFTHDSLIAPAGDYVTAKMG	60		
QY	61	IDEVIVRQNGSTRALPNCVRHRRGKTLVSYEAGNAGFVCSYNGMGFGSGELQSPFE	120		
DB	61	IDEVIVRQNGSTRALPNCVRHRRGKTLVNAEAGNAGFVCSYNGMGFGSGELQSPFE	120		
QY	121	KDLGESLNRKCLGLKEVARVESFHGFLYGCDFQDAPPPLMDYLDGDAWYLEPMPKHSGL	180		

```
Db 121 KELYGESLNKKCGLEKVEAVESFHGFIYGCFOEAPSLMDYLGDAWYLEPIFKHSGGL 180
OY 181 ELVGPCKVYIKANKKAPAEFNGDAVHVGWTHASSLSRSESTFSSLAGNAALPEPGAGL 240
Db 181 ELVGPCKVYIKANKKAPAEFNGDAVHVGWTHASSLSRSESTFSSLAGNAALPEPGAGL 240
OY 241 QMTSKYSGKGVLMIDYSGVSHADLVPBELMAFGAKOEKRLKEIGDVRARIYRSHLNCV 300
Db 241 QMTSKYSGKGVLMIDYSGVSHADLVPBELMAFGAKOEKRLKEIGDVRARIYRSHLNCV 300
OY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIVKEDMPEDLKRRLADSVQRTGPGFWMES 360
Db 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIVKEDMPEDLKRRLADSVQRTGPGFWMES 360
OY 361 DDNDNMTASQNKQKYSRSDLSNLGFGEDYDGYAVPGVYKSAIGETSYRGFYRAY 420
Db 361 DDNDNMTASQNKQKYSRSDLSNLGFGEDYDGYAVPGVYKSAIGETSYRGFYRAY 420
OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449
Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449
```

## RESULT 2

052142 PRELIMINARY: PRT: 449 AA.

AC 052142; 008194;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Iron-sulfur protein large subunit.

GN PAHAC.

OS Pseudomonas putida.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC NCBI\_TaxID=303;

RN NCBI\_TaxID=303;

RP SEQUENCE FROM N.A.

RC STRAIN=OUS82;

RA Takizawa N.;

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=OUS82;

RX MEDLINE=94209249; PubMed=8157615;

RA TAKIZAWA N., Kaide N., Torigoe S., Moritani T., Sawada T., Satoh S.,

RA Kiyohara H.; Identification and characterization of genes encoding polycyclic

RT aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon

RT dihydrodiol dehydrogenase in Pseudomonas putida OUS82.";

RL J. Bacteriol. 176:2444-2449(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=OUS82;

RA Noboru T., Yoshiya I., Takashi S., Kazuhiro Y., Masamichi K.,

RA Yue-Wu W., Masao F., Hozoh K.;

RT "The molecular analysis of NAH-type cluster located on the

RT chromosomes of Pseudomonas aeruginosa Paki and Pseudomonas putida

RT OUS82.";

RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.

RC EMBL; AB004059; BAA20391.1; -

DR HSSP; P23094; INDO.

DR InterPro: IPR001281; Rieseke.

DR InterPro: IPR001663; Ring\_hydroxyl\_A.

DR Pfam: PF00355; Rieseke.1.

DR Pfam: PF00848; Ring\_hydroxyl\_A.1.

DR PRINTS: PR00090; RINGDIOXGNASE.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA.1.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA.1.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA.1.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA.1.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA.1.

DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA.1.

Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

OY 1 MNYNKILVSESGLSQKHLIHGDELEFOHELKTIARNNLFLTHDSLIPAPGYVTAKMG 60

Db 1 MNYNKILVSESGLSQKHLIHGDELEFOHELKTIARNNLFLTHDSLIPAPGYVTAKMG 60

OY 61 IDEVIYSRONDSSIRAFVLCVCRHGRKTLVSEAGNAKGFVCSYHGMFGSNGELOSVPPE 120

Db 61 IDEVIYSRONDSSIRAFVLCVCRHGRKTLVSEAGNAKGFVCSYHGMFGSNGELOSVPPE 120

OY 121 KDLVSELNKKCGLEKVEAVESFHGFIYGCFOEAPSLMDYLGDAWYLEPIFKHSGGL 180

Db 121 KDLVSELNKKCGLEKVEAVESFHGFIYGCFOEAPSLMDYLGDAWYLEPIFKHSGGL 180

OY 181 ELVGPCKVYIKANKKAPAEFNGDAVHVGWTHASSLSRSESTFSSLAGNAALPEPGAGL 240

Db 181 ELVGPCKVYIKANKKAPAEFNGDAVHVGWTHASSLSRSESTFSSLAGNAALPEPGAGL 240

OY 241 QMTSKYSGKGVLMIDYSGVSHADLVPBELMAFGAKOEKRLKEIGDVRARIYRSHLNCV 300

Db 241 QMTSKYSGKGVLMIDYSGVSHADLVPBELMAFGAKOEKRLKEIGDVRARIYRSHLNCV 300

OY 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIVKEDMPEDLKRRLADSVQRTGPGFWMES 360

Db 301 FPNNSMLTCSGVFKVWNPIDANTTEWYTAIVKEDMPEDLKRRLADSVQRTGPGFWMES 360

OY 361 DDNDNMTASQNKQKYSRSDLSNLGFGEDYDGYAVPGVYKSAIGETSYRGFYRAY 420

Db 361 DDNDNMTASQNKQKYSRSDLSNLGFGEDYDGYAVPGVYKSAIGETSYRGFYRAY 420

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

OY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

Db 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449



```
Db 1 MNKNKTLVTESGLQKHLHGDEELFORELETFIARNNLFLTHDSLIPSPDYVTAKMG 60
QY 61 IDEVIVSRONDGSIRAFNLVNCRRHCKTLVSVAGNAKGFVCSYHGMGFGSNGELQSVPE 120
Db 61 VDEVIVSRONDGSIRAFNLVNCRRHCKTLVHAEGNAKGFVCSYHGMGFGANGLQSVPE 120
QY 121 KDLYGESLNKCKGLKEVARVESFHGFIYCGFDQAPPLMDYLGDAWYLEPMFKHSGGL 180
Db 121 KELYGALDKKCKGLKEVARVESFHGFIYCGFDQAPPLMDYLGDAWYLEPMFKHSGGL 180
QY 181 ELVGPFGKVIYIANKKAPAEFNVGDYAHVGMTHASLSRSGESIFSSLAGNALPPEGAGL 240
Db 181 ELIGPGKVIYIANKKAPAEFNVGDYAHVGMTHASLSRSGESIFSSLAGNALPPEGAGL 240
QY 241 QMTSKYSGSMGVLMGDSYGVSHADLVPELMAFGAKOEFLNKEIGDVRARIYRSHLNCV 300
Db 241 QMTSKYSGSMGVLMGDSYGVSHADLVPELMAFGAKOEFLNKEIGDVRARIYRSHLNCV 300
QY 301 FPNNSMLTCSGVFKYWNPIDANTTEWMTYAIYEKMPEDLKRRLDVSQRTIGPAGFWS 360
Db 301 FPNNSMLTCSGVFKYWNPIDANTTEWMTYAIYEKMPEDLKRRLDVSQRTIGPAGFWS 360
QY 361 DDNDMMETASQNGKKYQSSHSDLSNLGFGEDYVDYAVPGVYKSAIGETSYRGFTYR 420
Db 361 DDNDMMETESQNAKKYQSSHSDLSNLGFGEDYVDYAVPGVYKSAIGETSYRGFTYR 420
QY 421 QAHVSSNMWAEFEHASTWHTLELTKTDR 449
Db 421 GAHSSSNWAEFEHASTWHTLELTKTDR 449

RESULT 4
QY 0915D2: PRELIMINARY: PRT: 447 AA.
AC 0915D2:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Polyaromatic hydrocarbon dioxygenase large subunit.
GN PAHAC.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H;
RX MEDLINE=21306905; PubMed=11414329;
RA Moser R., Stahl U.;
RT "Insights into the genetic diversity of initial dioxygenases from PAH-
degrading bacteria.";
RL Appl. Microbiol. Biotechnol. 55:609-618(2001).
DR HMBL; AF252550; AAF72976.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rleske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rleske; 1.
DR PRINTS; PR00090; RINGDIIOXNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
SQ SEQUENCE 447 AA; 49548 MW; 2950EF36123A6F21 CRC64;

Query Match 91.1%; Score 2193; DB 2; Length 447;
Best Local Similarity 88.9%; Pred. No. 8,1e-161;
Matches 399; Conservative 27; Mismatches 21; Indels 2; Gaps 1;
```

```
Db 59 VDEVIVSRONDGSVRAFLNVCRRHCKTLVHAEGNAKGFVCSYHGMGFGSNGELQSVPE 118
QY 121 KDLYGESLNKCKGLKEVARVESFHGFIYCGFDQAPPLMDYLGDAWYLEPMFKHSGGL 180
Db 119 KELYGDALKKCKGLKEVPRISEFHGFIYCGFDQAPPLMDYLGDAWYLEPMFKHSGGL 178
QY 181 ELVGPFGKVIYIANKKAPAEFNVGDYAHVGMTHASLSRSGESIFSSLAGNALPPEGAGL 240
Db 179 ELVGPFGKVIYIANKKAPAEFNVGDYAHVGMTHASLSRSGESIFSSLAGNALPPEGAGL 238
QY 241 QMTSKYSGSMGVLMGDSYGVSHADLVPELMAFGAKOEFLNKEIGDVRARIYRSHLNCV 300
Db 239 QMTSKYSGSMGVLMGDSYGVSHADLVPELMAFGAKOEFLNKEIGDVRARIYRSHLNCV 298
QY 301 FPNNSMLTCSGVFKYWNPIDANTTEWMTYAIYEKMPEDLKRRLDVSQRTIGPAGFWS 360
Db 299 FPNNSMLTCSGVFKYWNPIDANTTEWMTYAIYEKMPEDLKRRLDVSQRTIGPAGFWS 358
QY 361 DDNDMMETASQNGKKYQSSHSDLSNLGFGEDYVDYAVPGVYKSAIGETSYRGFTYR 420
Db 359 DDNDMMETESQNAKKYQSSHSDLSNLGFGEDYVDYAVPGVYKSAIGETSYRGFTYR 418
QY 421 QAHVSSNMWAEFEHASTWHTLELTKTDR 449
Db 419 GAHSSSNWAEFEHASTWHTLELTKTDR 447

RESULT 5
QY 052382: PRELIMINARY: PRT: 447 AA.
AC 052382:
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Naphthalene 1,2 dioxygenase large oxygenase component.
GN NAGAC.
OS Ralstonia sp. U2.
OC Plasmid pMWU2.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=70356;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U2;
RX MEDLINE=98233751; PubMed=9573207;
RA Fuenmayor S.L., Wild M., Boyes A.L., Williams P.A.;
RT "A gene cluster encoding steps in conversion of naphthalene to
gentisate in Pseudomonas sp. strain U2.";
RL J. Bacteriol. 180:2522-2530(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=U2;
RX MEDLINE=20576173; PubMed=11133965;
RA Zhou N.Y., Fuenmayor S.L., Williams P.A.;
RT "nag genes of Ralstonia (formerly Pseudomonas) sp. strain U2 encoding
enzymes for gentisate catabolism.";
RL J. Bacteriol. 183:700-708(2001).
DR EMBL; AF036940; AADI2610.1; -.
DR HSSP; P23094; INDO.
DR InterPro; IPR001281; Rleske.
DR InterPro; IPR001663; Ring_hydroxyl_A.
DR Pfam; PF00355; Rleske; 1.
DR PRINTS; PR00090; RINGDIIOXNASE.
DR PROSITE; PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase; Plasmid.
SQ SEQUENCE 447 AA; 49570 MW; 4553AAF4B4410ED0 CRC64;

Query Match 90.9%; Score 2189; DB 2; Length 447;
Best Local Similarity 89.1%; Pred. No. 1,6e-160;
Matches 400; Conservative 25; Mismatches 22; Indels 2; Gaps 1;
```

```
Db 1 MIYEN--LVEEAGLTQKHLIHGDELFQHELTIFARNWMLFTHDSLIPSPGDYVYAKMG 58
QY 61 IDEVIYSRQNDGSTRALFNLNCRHGRKTLVSEAGNAGKGFVCSYHGNGFSGNGLQSVPE 120
Db 59 VDEVIYSRQNDGSTRALFNLNCRHGRKTLVSEAGNAGKGFVCSYHGNGFSGNGLQSVPE 118
QY 121 KDLGSELNKCKLGLKEVARSEFHGFIYGCFOEAPPLMDYLGDAWYLEPMEKHSGL 180
Db 119 KELYGDIKKKKCLGLKEVPRIESHFGIYGCFOEAPPLMDYLGDAWYLEPMEKHSGL 178
QY 181 ELVPPGKVVYIKANWKPAENFVGDAYHVGWTHASLSRSGESIFSSLAGNALPPEGAGL 240
Db 179 ELVPPGKVVYIKANWKPAENFVGDAYHVGWTHASLSRSGESIFSSLAGNALPPEGAGL 238
QY 241 QMTSKYSGGVLMDYSGVHSADLVPDLMAFGAKOEKRLKEIGDVARARYSHLNTV 300
Db 239 QMTSKYSGGVLMDYSGVHSADLVPDLMAFGAKOEKRLKEIGDVARARYSHLNTV 298
QY 301 FPNNSMLTCSGVFKNPDPIDANTTEVNTYAIVEKMPEDLKRRLADSVORTIGPAGFWS 360
Db 299 FPNNSMLTCSGVFKNPDPIDANTTEVNTYAIVEKMPEDLKRRLADSVORTIGPAGFWS 358
QY 361 DDNDNMTASONGKKYOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
Db 359 DDNDNMTASONGKKYOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 418
QY 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449
Db 419 QAHVSSNMAEFHASSTWHTELTKTTDR 447
```

RESULT 6  
08VUD4 PRELIMINARY: PRT: 447 AA.

```
AC 08VUD4;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE DataC.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group.
OC Burkholderia.
OX NCBI_TaxID=292;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-R34;
RX MEDLINE=20254695; PubMed=10795678;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Properties of the trihydroxytoluene oxygenase from Burkholderia
   cepacia R34: an extradiol dioxygenase from the 2,4-dinitrotoluene
   pathway."
RT Arch. Microbiol. 173:86-90(2000).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN-R34;
RA Johnson G.R., Jain R.K., Spain J.C.;
RT "Origins of the 2,4-dinitrotoluene pathway."
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF169302; AAL50021.1; -.
DR InterPro: IPR001281; Rieseke.
DR Pfam: PF00355; Rieseke.
DR Pfam: PF00848; Ring_hydroxyl_A.
DR PRINTS: PR00090; RINGDIPOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; UNKNOWN_1.
SQ SEQUENCE 447 AA: 49468 MW: 2458583320275C CRC64;
```

Query Match 86.1%; Score 2074; DB 2; Length 447;  
Best Local Similarity 83.5%; Pred. No. 12e-151;  
Matches 375; Conservative 39; Mismatches 33; Indels 2; Gaps 1;  
QY 1 MNYNKILVESGLISOKHLIHGDELFQHELTIFARNWMLFTHDSLIPAGDYVYAKMG 60

```
Db 1 MSYGN--LVEEAGLTQKHLIHGDELFQHELTIFARNWMLFTHDSLIPSPGDYVYAKMG 58
QY 61 IDEVIYSRQNDGSTRALFNLNCRHGRKTLVSEAGNAGKGFVCSYHGNGFSGNGLQSVPE 120
Db 59 VDEVIYSRQNDGSTRALFNLNCRHGRKTLVSEAGNAGKGFVCSYHGNGFSGNGLQSVPE 118
QY 121 KDLGSELNKCKLGLKEVARSEFHGFIYGCFOEAPPLMDYLGDAWYLEPMEKHSGL 180
Db 119 KELYGDAIKKKCLGLKEVPRIESHFGIYGCFOEAPPLMDYLGDAWYLEPMEKHSGL 178
QY 181 ELVPPGKVVYIKANWKPAENFVGDAYHVGWTHASLSRSGESIFSSLAGNALPPEGAGL 240
Db 179 ELVPPGKVVYIKANWKPAENFVGDAYHVGWTHASLSRSGESIFSSLAGNALPPEGAGL 238
QY 241 QMTSKYSGGVLMDYSGVHSADLVPDLMAFGAKOEKRLKEIGDVARARYSHLNTV 300
Db 239 QMTSKYSGGVLMDYSGVHSADLVPDLMAFGAKOEKRLKEIGDVARARYSHLNTV 298
QY 301 FPNNSMLTCSGVFKNPDPIDANTTEVNTYAIVEKMPEDLKRRLADSVORTIGPAGFWS 360
Db 299 FPNNSMLTCSGVFKNPDPIDANTTEVNTYAIVEKMPEDLKRRLADSVORTIGPAGFWS 358
QY 361 DDNDNMTASONGKKYOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 420
Db 359 DDNDNMTASONGKKYOSRSDLSNLGFGEDYGDVAVPGVYKSAIGETSYRGFYRAY 418
QY 421 QAHVSSNMAEFHASSTWHTELTKTTDR 449
Db 419 QAHVSSNMAEFHASSTWHTELTKTTDR 447
```

RESULT 7  
P95564 PRELIMINARY: PRT: 447 AA.

```
AC P95564;
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ISPalpa 2MT.
GN NMDAC.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-JS42;
RX MEDLINE=97128768; PubMed=8973308;
RA Parales J.V., Kumar A., Parales R.E., Gibson D.T.;
RT "Cloning and sequencing of the genes encoding 2-nitrotoluene
   dioxygenase from Pseudomonas sp. JS42."
RT Gene 181:57-61(1996).
RL Gene 181:57-61(1996).
DR EMBL: U49504; AAB40383.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR Pfam: PF00355; Rieseke.
DR Pfam: PF00848; Ring_hydroxyl_A.
DR PRINTS: PR00090; RINGDIPOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
SQ SEQUENCE 447 AA: 49485 MW: 1CB0E223E528E3BD CRC64;
```

Query Match 85.4%; Score 2056; DB 2; Length 447;  
Best Local Similarity 84.0%; Pred. No. 3e-150;  
Matches 377; Conservative 34; Mismatches 36; Indels 2; Gaps 1;  
QY 1 MNYNKILVESGLISOKHLIHGDELFQHELTIFARNWMLFTHDSLIPAGDYVYAKMG 60  
Db 1 MSYGN--LVEEAGLTQKHLIHGDELFQHELTIFARNWMLFTHDSLIPSPGDYVYAKMG 58  
QY 61 IDEVIYSRQNDGSTRALFNLNCRHGRKTLVSEAGNAGKGFVCSYHGNGFSGNGLQSVPE 120  
Db 59 VDEVIYSRQNDGSTRALFNLNCRHGRKTLVSEAGNAGKGFVCSYHGNGFSGNGLQSVPE 118

QY	121	KDJVGEJSLNKKCLCLGKEVARESHESHGITGCGFQDEAPPLMDYIGDAWYIEPFFKSGGL	180
Db	119	KELVDGADKKKCLGKEVARESHESHGITGCGFQDEAPPLMDYIGDAWYIEPFFKSGGL	178
QY	181	ELVGPBGKVVYIKAMWKAPEAEFNVDAHVHVMTHASSLSRGEISFSSLAGMAALPPEGAGI	240
Db	179	ELVGPBGKVVYIKAMWKAPEAEFNVGDIVHGMTHAALRAGOSVFSSLAGMAALPPEGAGI	238
QY	241	QMTSKYSGGMVLMDYVSGVHSADLYPELMAFGAKOEBLKEIGVRAITYSHLNCIV	300
Db	239	QMTSKYSGGMVLMDYVSGVHSADLYPELMAFGAKOEBLKEIGVRAITYSHLNGIV	298
QY	301	FPNNSMLTCSGVFKVWNPIDANTIEVVTYIAIVEKMDIEDLRRLADSVORTIIPAGFWS	360
Db	299	FPNNSFLTGSAITFKVWNPIDENTIEVVTYIAIVEKMDIEDLRRLADQAORSIIPAGFWS	358
QY	361	DDNDIMETASONGKRGYSDBDLSSNGPEPDYGAUWYGVVYKSAIGSTYRGRGYRAY	420
Db	359	DDNEMMELSONMAKRYOSSNSDOJASISGFKDYGDECTGVGVYKSAIGSTYRGRYRAY	418
QY	421	QAHVSSSNMAAEFEHASTWHTELTKTTDR	449
Db	419	QAHVSSSNMAAEFEHASTWHTELTKTTDR	447

RESULT	8
ID	Q8RTL4
AC	Q8RTL4
DT	01-JUN-2002 (TREMBLER, 21, Created)
DT	01-JUN-2002 (TREMBLER, 21, Last sequence update)
DT	01-JUN-2002 (TREMBLER, 21, Last annotation update)
DE	Oxygenase-alpha NBDP.
GN	NB2AC.
OS	Comamonas sp., JS765.
OC	Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas
OX	NCBI_TaxID=58226;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=JS765;
RA	MEDLINE=11681044; PubMed=11823201;
RT	Lesener D.J., Johnson G.R., Parales R.E., Spain J.C., Gibson D.T.;
RT	"Molecular Characterization and Substrate Specificity of Nitrobenzene
RL	Dioxygenase from Comamonas sp. Strain JS765.";
RL	Appl. Environ. Microbiol. 68:624-641(2002).
SO	EMBL; AF379638; AAL76202.1; -
SO	SEQUENCE 447 AA; 49556 MW; 4CA961F99A37DF3A CRC64;

Query Match	84.28;	Score 2027;	DB 2;	Length 447;
Best Local Similarity	82.28;	Pred. No. 5.1e-148;		
Matches 369; Conservative	40;	Mismatches 38;	Indels 2;	Gaps 1;

Db	1	MSYON--LVSPAGLTQVLLIHGDKELEFQHEHLKLTIPARINMLFETHDSLITSPGDYAAAKG	58
Qy	61	IDEYIVSRQNDGSTRAFLNANCRRHGRKTLVSVEAGNAKGFVCSYHGMGFGSGNGLQSVPEE	120
Db	59	VDEYIVSRQNDGSRALANCRHGRGKTLVIAEAGNAKGFVCSYHGMGFGSGNGLQSVPEE	118
Qy	121	KDLGESLNKKCLGLGKEAVARVESHGFIYGCFOEAPPLMDLYGDAAMWLEPFKFKSGTL	180
Db	119	KELYGDAIRKKKCLGLGKEVPRIESPFGFIYGCFOEAPPLIDLYGDAAMWLEPFKFKSGTL	178
Qy	181	ELVGPCKVYVYKAMMKRAPNPFYGDAYHVGWTHAASLSRSESTFSSLAGNALPPRGAGL	240
Db	179	ELVGPCKVYVYKAMMKSFALNPFYGDGIHVGWTHAALRQGVYFSSIAENAKLPPRGAGL	238
Qy	241	QMTSKYSGGMGLVMDGYSGVSHADLPELTAFGAQRINKTIGVVARIRYSHLNCYV	300
Db	239	QMTSKYSGGMGLVYSGVSGNFSADMDLIDLAFGAQRKAKTIGVVARIRYSPFLNGTI	298

[illegible]

```

RESULT 9
ID      Q45695      PRELIMINARY;      PRT;      451 AA.
AC      Q45695;
DT      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      ISP-alpha.
GN      DNTAC.
OS      Burkholderia sp. (strain RASC).
OC      Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC      Burkholderia
OX      NCBI_TaxID=69003;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-DNT;
RX      MEDLINE=96326338; Pubmed=8759857;
RA      Suen W.C., Haigler B.E., Spain J.C.;
RT      "2,4-Dinitrocoluene dioxygenase from Burkholderia sp. strain DNT:
RT      similarity to naphthalene dioxygenase.";
RL      J. Bacteriol. 178:4926-4934(1996).
DR      EMBL: U62430; AAB09766.1; -.
DR      HSSP; P23094; INDO.
DR      InterPro; IPR001281; Rieske.
DR      InterPro; IPR001663; Ring_hydroxyl_A.
DR      Pfam; PF00355; Rieske; 1.
DR      Pfam; PF00848; Ring_hydroxyl_A; 1.
DR      PRINTS; PRO0090; RINGDIOXGNASE.
DR      PROSITE; PS00570; RING-HYDROXYL_ALPHA; 1.
SQ      SEQUENCE 451 AA; 49827 MW; 726796C29CE9A10 CRC64;

```

Query Match	81.7%	Score 1967.5;	DB 2;	Length 451;
Best Local Similarity	80.0%	Pred. NO. 2e-143;		
Matches 359; Conservative	41;	Mismatches 46;	Indels 3;	Gaps 2;

QY	1	MNI	NK	IL	VSS	SG	SQ	KH	LH	IG	DE	EL	Q	H	E	L	T	F	A	N	N	L	F	L	H	D	S	L	I	P	A	P	G	O	Y	T	A	K	K	60															
Db	6	M	S	Y	O	N	-	L	V	S	E	L	T	O	K	H	L	L	Y	G	D	E	L	F	O	H	E	L	K	T	F	A	N	N	L	F	L	H	D	S	L	I	P	S	P	G	O	Y	T	A	K	K	63		
QY	61	I	D	E	V	I	S	H	O	N	G	S	T	R	A	F	L	N	C	R	H	G	K	T	L	V	S	E	A	G	M	A	G	F	C	S	H	G	M	G	F	O	S	N	G	L	O	S	V	P	E	120			
Db	64	V	D	E	V	I	S	H	O	N	G	S	T	R	A	F	L	N	C	R	H	G	K	T	L	V	S	E	A	G	M	A	G	F	C	S	H	G	M	G	F	O	S	N	G	L	O	S	V	P	E	123			
QY	121	K	D	L	E	G	S	L	N	K	K	C	L	G	L	E	V	A	R	E	S	H	G	F	I	Y	G	C	F	D	O	E	A	P	L	M	D	I	G	D	A	M	L	E	P	M	F	K	H	S	G	L	180		
Db	124	K	E	L	V	G	D	A	L	K	K	C	L	G	L	E	V	P	R	E	S	H	G	F	I	Y	G	C	F	D	O	E	A	P	L	M	D	I	G	D	A	M	L	E	P	M	F	K	H	S	G	L	183		
QY	181	E	L	V	G	P	E	G	K	V	I	T	A	K	M	K	P	A	E	N	F	G	D	A	I	H	V	G	M	T	H	A	S	S	R	S	E	S	I	F	S	S	L	A	G	N	A	L	P	R	E	G	A	L	240
Db	184	E	L	V	G	P	A	V	V	V	K	G	N	M	K	F	A	E	N	F	V	D	I	H	I	G	M	T	H	A	S	I	R	A	G	A	I	F	A	P	L	A	G	N	A	M	L	P	R	E	G	T	243		
QY	241	O	M	T	S	K	Y	S	G	M	V	L	M	D	G	S	V	C	H	S	A	D	I	V	E	L	M	A	F	G	A	K	E	R	L	K	E	I	G	V	R	A	I	T	Y	S	H	L	A	N	C	I	Y	300	
Db	244	Q	A	T	T	K	Y	G	S	G	I	S	V	L	D	A	S	V	G	A	D	I	V	E	M	A	F	G	A	K	E	R	L	K	E	I	G	V	R	A	I	T	Y	S	O	N	G	E	V	Y	303				
QY	301	F	P	N	N	S	M	L	T	C	S	G	E	V	F	V	N	M	P	I	A	N	T	E	V	T	A	I	V	E	K	D	M	E	D	L	K	R	L	A	D	O	S	V	T	O	G	P	A	F	E	W	S	360	
Db	304	F	P	N	N	C	F	L	T	G	A	V	E	V	F	N	P	I	D	E	N	T	E	A	M	T	A	I	V	E	K	D	M	E	D	L	K	R	L	A	D	A	O	A	S	T	O	P	A	I	T	W	S	363	
QY	361	D	D	N	D	N	M	E	T	A	S	O	N	G	K	K	Y	S	R	D	S	L	N	G	F	E	D	Y	G	D	A	V	I	P	G	V	V	K	S	A	I	G	T	S	R	G	E	R	A	Y	420				

```
Db 364 DDNNM-VLSGNARKYSSNSDLADLGFGRDVGDECPGVSSKSAFSETHNRGFYAY 422
QY 421 CAHYSSNMWEEFHAASSFWHETRTTDR 449
Db 423 CAHSSNMWEEFENTSRMWHETRTTDR 451

RESULT 10
Q9ETK2 PRELIMINARY; PRT: 277 AA.
ID 09ETK2;
AC 09ETK2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2 OR NAHAC.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2JIDNH, PRIMN1, AND 31A2NH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF306438; AAC25698.1; -.
DR EMBL: AF306432; AAC25692.1; -.
DR EMBL: AF306436; AAC25696.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIIOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 277 AA; 30094 MW; BECF9B5A379546DB CRC64;

Query Match 61.9%; Score 1490; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 6.4e-107;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 56 TAKMGIDEIVYSRONDSIRAFNLVCRHKGKTLVSVENGNAKGFVCSYHGMFGSNGELQ 115
Db 1 TAKMGIDEIVYSRONDSIRAFNLVCRHKGKTLVSVENGNAKGFVCSYHGMFGSNGELQ 60
QY 116 SVPEKDLVGSLSNKKCLGKEVARVESFHGFIYGCPEDEAPPLMDYLGDAWYLEPMFK 175
Db 61 SVPEKDLVGSLSNKKCLGKEVARVESFHGFIYGCPEDEAPPLMDYLGDAWYLEPMFK 120
QY 176 HSGGLELVGPPGKVYIKANKKAPAEENFVGDAYHWGTHASSLSRGSSEIFSSLAGNAALPP 235
Db 121 HSGGLELVGPPGKVYIKANKKAPAEENFVGDAYHWGTHASSLSRGSSEIFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPELMAFGAKOERLNKEIGDVARIRYRSH 295
Db 181 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPELMAFGAKOERLNKEIGDVARIRYRSH 240
QY 296 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 332
Db 241 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 277

RESULT 11
Q9FS54 PRELIMINARY; PRT: 277 AA.
ID 09FS54;
AC 09FS54;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
```

```
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein 2 (Fragment).
GN NAHAC2.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=511ASAL;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF306440; AAC25700.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIIOXNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Dioxygenase.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 277 AA; 30094 MW; FADFDA5D37E241AB CRC64;

Query Match 61.8%; Score 1486; DB 2; Length 277;
Best Local Similarity 99.6%; Pred. No. 9.1e-107;
Matches 276; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 56 TAKMGIDEIVYSRONDSIRAFNLVCRHKGKTLVSVENGNAKGFVCSYHGMFGSNGELQ 115
Db 1 TAKMGIDEIVYSRONDSIRAFNLVCRHKGKTLVSVENGNAKGFVCSYHGMFGSNGELQ 60
QY 116 SVPEKDLVGSLSNKKCLGKEVARVESFHGFIYGCPEDEAPPLMDYLGDAWYLEPMFK 175
Db 61 SVPEKDLVGSLSNKKCLGKEVARVESFHGFIYGCPEDEAPPLMDYLGDAWYLEPMFK 120
QY 176 HSGGLELVGPPGKVYIKANKKAPAEENFVGDAYHWGTHASSLSRGSSEIFSSLAGNAALPP 235
Db 121 HSGGLELVGPPGKVYIKANKKAPAEENFVGDAYHWGTHASSLSRGSSEIFSSLAGNAALPP 180
QY 236 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPELMAFGAKOERLNKEIGDVARIRYRSH 295
Db 181 EGAGLQMTSKYSGSMGVLMDCYSGVSHSADLVPELMAFGAKOERLNKEIGDVARIRYRSH 240
QY 296 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 332
Db 241 LNCITVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 277

RESULT 12
Q9FS73 PRELIMINARY; PRT: 277 AA.
ID 09FS73;
AC 09FS73;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Naphthalene dioxygenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 191IDNH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139738;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=191IDNH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxygenase genes (nahac) in Pseudomonas strains from West Mediterranean Sea.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL: AF306426; AAG25686.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KM Dioxigenase.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 30064 MW; DB47868EB6DD525A CRC64;

Query Match
Best Local Similarity 61.7%; Score 1486; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRQNDGSTRAPFLNVCRRHGRKTLVSVEAGNAGFVCSYHGWMFGSGNGELQ 115
DB 1 TAKMGIDEIVSRQNDGSTRAPFLNVCRRHGRKTLVSVEAGNAGFVCSYHGWMFGSGNGELQ 60

QY 116 SVPEEKDLVSGESLNKKCLGKEVARVESFHGFIYGCFFDQDEAPPLMDYLGDAWYLEPMPK 175
DB 61 GVPEEKDLVSGESLNKKCLGKEVARVESFHGFIYGCFFDQDEAPPLMDYLGDAWYLEPMPK 120

QY 176 HSGGLELVGPPGKVVIKANWKAPAENFVGDAYHVGMTTHASSLRSGESIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVIKANWKAPAENFVGDAYHVGMTTHASSLRSGESIFSSLAGNAALPP 180

QY 236 EGAGLQMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 295
DB 181 EGAGLQMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 240

QY 296 LNCVTFPPNNSMLTCSGVFFKVMNPIDANTTEVWTYAIY 332
DB 241 LNCVTFPPNNSMLTCSGVFFKVMNPIDANTTEVWTYAIY 277

RESULT 13
Q9F5S8 PRELIMINARY: PRT: 277 AA.
ID 09F5S8;
AC Q9F5S8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Naphthalene dioxigenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. PR3MN2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PR3MN2;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxigenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RT Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF306434; AAG25694.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KM Dioxigenase.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 30064 MW; 6BFF83D32F04CFC3 CRC64;

Query Match
Best Local Similarity 99.6%; Score 1484; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRQNDGSTRAPFLNVCRRHGRKTLVSVEAGNAGFVCSYHGWMFGSGNGELQ 115
DB 1 TAKMGIDEIVSRQNDGSTRAPFLNVCRRHGRKTLVSVEAGNAGFVCSYHGWMFGSGNGELQ 60

QY 116 SVPEEKDLVSGESLNKKCLGKEVARVESFHGFIYGCFFDQDEAPPLMDYLGDAWYLEPMPK 175
DB 61 SVPEEKDLVSGESLNKKCLGKEVARVESFHGFIYGCFFDQDEAPPLMDYLGDAWYLEPMPK 120

QY 176 HSGGLELVGPPGKVVIKANWKAPAENFVGDAYHVGMTTHASSLRSGESIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVIKANWKAPAENFVGDAYHVGMTTHASSLRSGESIFSSLAGNAALPP 180

QY 236 EGAGLQMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 295
DB 181 EGAGLQMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 240

QY 296 LNCVTFPPNNSMLTCSGVFFKVMNPIDANTTEVWTYAIY 332
DB 241 LNCVTFPPNNSMLTCSGVFFKVMNPIDANTTEVWTYAIY 277

RESULT 14
Q9F5S9 PRELIMINARY: PRT: 277 AA.
ID 09F5S9;
AC Q9F5S9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Naphthalene dioxigenase iron sulfur protein (Fragment).
GN NAHAC.
OS Pseudomonas sp. 8IDINH.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=139736;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8IDINH;
RA Ferrero M.A., Lalucat J., Bosch R.;
RT "Coexistence of two naphthalene dioxigenase genes (nahac) in
Pseudomonas strains from West Mediterranean Sea.";
RT Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF306433; AAG25693.1; -.
DR HSSP: P23094; INDO.
DR InterPro: IPR001281; Rieseke.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieseke; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KM Dioxigenase.
FT NON_TER 1 1
FT NON_TER 277 277
SQ SEQUENCE 277 AA; 29995 MW; AAD9A9B5A378F6DB CRC64;

Query Match
Best Local Similarity 61.6%; Score 1483; DB 2; Length 277;
Matches 276; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 TAKMGIDEIVSRQNDGSTRAPFLNVCRRHGRKTLVSVEAGNAGFVCSYHGWMFGSGNGELQ 115
DB 1 TAKMGIDEIVSRQNDGSTRAPFLNVCRRHGRKTLVSVEAGNAGFVCSYHGWMFGSGNGELQ 60

QY 116 SVPEEKDLVSGESLNKKCLGKEVARVESFHGFIYGCFFDQDEAPPLMDYLGDAWYLEPMPK 175
DB 61 SVPEEKDLVSGESLNKKCLGKEVARVESFHGFIYGCFFDQDEAPPLMDYLGDAWYLEPMPK 120

QY 176 HSGGLELVGPPGKVVIKANWKAPAENFVGDAYHVGMTTHASSLRSGESIFSSLAGNAALPP 235
DB 121 HSGGLELVGPPGKVVIKANWKAPAENFVGDAYHVGMTTHASSLRSGESIFSSLAGNAALPP 180

QY 236 EGAGLQMTSKYSGMGVLDGYSGVHSADLVPELMAFGAKQERLKEIGDVARIRYRSH 295
```

Db 181 EGAGLQMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKQEGLNKEIGDVRARIYRSH 240  
QY 296 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 332  
Db 241 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 277

## RESULT 15

Q9F5S3 PRELIMINARY; PRT; 277 AA.  
AC Q9F5S3;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Naphthalene dioxygenase iron sulfur protein 1 (Fragment).  
GN NAHAC1.  
OS Pseudomonas putida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=511ANH; Talucat J., Bosch R.;  
RA Ferrero M.A., of two naphthalene dioxygenase genes (nahAc) in  
RT "Coexistence of two naphthalene dioxygenase genes (nahAc) in  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF306441; AAC25701.1; "  
DR HSSP: P23094; 1NDO.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001653; Ring\_hydroxyl\_A.  
DR Pfam: PF00355; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS: PR00090; RINGDIOXGNASE.  
DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
KW Dioxygenase.  
KW NON\_TER 1  
FT 277  
FT NON\_TER 1  
SQ SEQUENCE 277 AA; 30062 MW; 82C78D865A92D0AB CRC64;

## Query Match

Best Local Similarity 60.3%; Score 1451; DB 2; Length 277;

Matches 266; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 56 TAKMGIDEVIYSRONDSIRAFILNVCRRHGTIVSVEAGNNAKGFVCSYHGNGFGSGNGELQ 115  
Db 1 TAKMGIDEVIYSRQSDSIRAFILNVCRRHGTIVNNAEAGNAKGFVCSYHGNGFGSGNGELQ 60  
QY 116 SVPEKDLYGESLNKKCLGLKEVARVESFHGFIYGCDFQDEAPPLMDYLGDAAWYLEPMFK 175  
Db 61 SVPEKELYGESLNKKCLGLKEVARVESFHGFIYGCDFQDEAPSLMDYLGDAAWYLEPIFK 120  
QY 176 HSGGLELVGPPGKVIKANKAPAEENFVGDAYHVGWTHASSLRSGESTFSSLAGNAALPP 235  
Db 121 HSGGLELVGPPGKVIKANKAPAEENFVGDAYHVGWTHASSLRSGESTFSSLAGNAALPP 180  
QY 236 EGAGLQMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKQERLNKEIGDVRARIYRSH 295  
Db 181 EGAGLQMTSKYSGMGVLMDSYSGVHSADLVPELMAFGAKQERLNKEIGDVRARIYRSH 240  
QY 296 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 332  
Db 241 LNCVFPNNSMLTCSGVFKVWNPIDANTTEVWYTAIV 277

Search completed: January 27, 2003, 09:02:25  
Job time : 27.8333 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:47:32 ; Search time 7.16667 seconds  
(without alignments)  
2598.540 Million cell updates/sec

Title: US-09-843-250-36

Sequence: 1 MNYNKILVSSGSLSQKHLI.....AEFEHASTWHTLTKTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2404	99.8	449	1 NDOB_PSEPU	P23094 pseudomonas
2	2318	96.3	449	1 NDOB_PSEFL	O07824 pseudomonas
3	2219	92.2	449	1 NDOB_PSEAE	O51494 pseudomonas
4	737	30.6	453	1 HCAE_ECOLI	O47139 escherichia
5	652.5	27.1	450	1 BED1_PSEPU	O07944 pseudomonas
6	650.5	27.0	450	1 TOD1_PSEPU	P13450 pseudomonas
7	617.5	25.6	458	1 BPHA_BURCE	P37333 butkholderi
8	617	25.6	457	1 BPHA_COMTE	O46372 comamonas t
9	617	25.6	458	1 BPHA_PSEPS	P08084 pseudomonas
10	588.5	24.4	448	1 BNA_PSESL	O52438 pseudomonas
11	587	24.4	458	1 BPA1_PSESL	P05438 pseudomonas
12	410.5	17.0	454	1 XYLX_PSEPU	P23099 pseudomonas
13	399	16.6	461	1 BENA_PSEPU	P07769 pseudomonas
14	249	10.3	374	1 YEAW_ECOLI	P76253 acinetobact
15	239.5	9.9	442	1 CHMO_AMATR	O93xex1 amaranthus
16	233	9.7	438	1 CHMO_ATRHO	O91xex1 atriplex ho
17	224	9.3	446	1 CHMO_BETVU	O22553 beta vulgar
18	223	9.3	439	1 CHMO_SPIOL	O04121 spinacia ol
19	216	9.0	422	1 CHMO_ARATH	O95xer0 arabidopsis
20	163	6.8	439	1 PHT3_PSEPU	O05183 pseudomonas
21	118.5	4.9	354	1 VANA_PSESP	O05616 pseudomonas
22	111	4.6	1276	1 PMF6_CHLNP	O94899 chlamydia p
23	110	4.6	329	1 VANA_PSES9	P12609 pseudomonas
24	110	4.6	432	1 CABA_COMTE	O44256 comamonas t
25	102	4.2	409	1 POBA_PSEPS	O52185 pseudomonas
26	101.5	4.2	1926	1 LPH_RABIT	P09849 oryctolagus
27	100.5	4.2	468	1 PPAL_PICPA	P52291 plicha past
28	99.5	4.2	3255	1 POLG_LMVE	P88876 1 genome po
29	99.5	4.1	543	1 7UP1_DROME	P16375 drosophila
30	99.5	4.1	746	1 7UP2_DROME	P16376 drosophila
31	98.5	4.1	1411	1 Y297_HUMAN	O10400 homo sapien
32	96.5	4.0	3255	1 POLG_LMVO	P31999 1 genome po
33	95	3.9	917	1 STL_STRAU	P41972 staphylococ

34	94.5	3.9	331	1 LDHA_RHIDE	O9PW58 rhizophila
35	94.5	3.9	350	1 UNRI_HUMAN	O9Y3F4 homo sapien
36	94.5	3.9	452	1 F26_YEAST	P32604 saccharomyc
37	94.5	3.9	518	1 ATPA_ENTHR	P26679 enterococcu
38	94	3.9	847	1 ORP8_HUMAN	O9B2F1 homo sapien
39	93	3.9	420	1 DHE3_PYRO	O59650 pyrococcus
40	93	3.9	1693	1 POLN_HEVMY	O04610 hepatitis e
41	92	3.8	985	1 AGIU_ASPOR	O12558 aspergillus
42	91.5	3.8	428	1 GRAL_CANAL	P28868 candida alb
43	91.5	3.8	849	1 DG3_MOUSE	P70175 mus musculu
44	91.5	3.8	849	1 DG3_RAT	O62936 rattus norv
45	91	3.8	405	1 DCP2_PEA	P51851 pisum sativ

## ALIGNMENTS

RESULT 1  
ID NDOB\_PSEPU STANDARD; PRT: 449 AA.  
AC P23094: 052124: 03461: 007830:  
DT 01-NOV-1991 (rel. 20, created)  
DT 01-NOV-1991 (rel. 20, last sequence update)  
DT 15-JUN-2002 (rel. 41, last annotation update)  
DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene  
DE 1,2-dioxygenase ISP alpha).  
GN NDOB OR NAHAC OR DOXB OR NAHA3 OR NDOC2.  
OS Pseudomonas putida, and  
OS Pseudomonas sp. (strain C18).  
OC Plasmid pDRI1, Plasmid NAH7, and Plasmid NP1.  
CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
CC Pseudomonas.  
OX NCBI\_TaxID=303, 306;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=89211973; PubMed=324348;  
RA Kurkela S., Leivaeslahti H., Palva E.T., Teeri T.H.;  
RT "Cloning, nucleotide sequence and characterization of genes encoding  
RT naphthalene dioxygenase of Pseudomonas putida strain NCIB9816.";  
RL Gene 73:355-362(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=NCIB 9816;  
RX MEDLINE=94042852; PubMed=8226631;  
RA Demone S.A., Stanley D.C., Olson E.S., Young K.D.;  
RT "Metabolism of dibenzothiole and naphthalene in Pseudomonas  
RT strains: complete DNA sequence of an upper naphthalene catabolic  
RT pathway.";  
RL J. Bacteriol. 175:6890-6901(1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=C7 / ATCC 17485; PLASMID=NAH7;  
RX MEDLINE=93252277; PubMed=8486285;  
RA Simon M.J., Ossling T.D., Saunders R., Ensley B.D., Suggs S.,  
RA Harcourt A.A., Suen W.-C., Cruden D.L., Gibson D.T., Zylstra G.J.;  
RT "Sequences of genes encoding naphthalene dioxygenase in Pseudomonas  
RT putida strains C7 and NCIB 9816-4.";  
RL Gene 127:31-37(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC SPECIES=P.putida; STRAIN=BS202; PLASMID=NP1;  
RA Bezdorodnikov S.G., Boronin A.M., Tiedje J.M.;  
RT "Nucleotide sequences of genes encoding an upper pathway of  
RT naphthalene metabolism of NP1 plasmid from Pseudomonas putida strain  
RT BS202.";  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
RN [6]

RP SEQUENCE FROM N.A.  
 RC SPECIES=P.putida; STRAIN=ATCC 17484;  
 RA Hamann C.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
 RC SPECIES=P.putida; STRAIN=NCIB 9816;  
 RX MEDLINE=98298434; PubMed=9634695;  
 RA Knappl B., Lee K., Carredano E., Parales R.E., Gibson D.T., Eklund H.,  
 Ramaswamy S.;  
 RT "Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene  
 1,2-dioxygenase";  
 RL Structure 6:571-586(1998).  
 CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
 CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
 CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
 CC NAPHTHALENE DIHYDRODIOL.  
 CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) -> (1R,2S)-1,2-  
 CC dihydronaphthalene-1,2-diol + NAD(+).  
 CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
 CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES  
 CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF  
 CC DIBENZOHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO  
 CC 1-HYDROXY-2-NAPHTHOIC ACID AND THE METABOLISM OF DBT IS LIMITED TO  
 CC OXIDATION OF THE AROMATIC RING.  
 CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME  
 CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON  
 CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED  
 CC OF FERREDOXIN REDUCTASE (NDOX) AND FERREDOXIN (NDOA) AND ISP IS  
 CC COMPOSED OF AN HEXAMER OF THREE LARGE ALPHA SUBUNITS (NDOB) AND  
 CC THREE SMALL BETA SUBUNITS (NDOC).  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 CC DIOXYGENASE ALPHA SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M23914; AAB47591.1; -;  
 DR EMBL: U49496; AAA92141.1; -;  
 DR EMBL: M60405; AAA16125.1; -;  
 DR EMBL: M83949; AAA25902.1; -;  
 DR EMBL: AF010471; AAB62707.1; -;  
 DR EMBL: AF004284; AAB61373.1; -;  
 DR PIR: JS0071; JS0071.  
 DR PIR: B49343; B49343.  
 DR PDB: INDO; 23-MAR-99.  
 DR InterPro: IPR001281; Rleske.  
 DR InterPro: IPR001663; Rling\_hydroxyl\_A.  
 DR Pfam: PF00335; Rleske; 1.  
 DR Pfam: PF00848; Rling\_hydroxyl\_A; 1.  
 DR PRINTS: PR00090; RINGDIOXGNASE.  
 DR PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.  
 DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
 DR Dioxygenase; NAD; Plasmid; 3D-structure.  
 KW METAL 81 81 IRON-SULFUR (2FE-2S).  
 FT METAL 83 83 IRON-SULFUR (2FE-2S).  
 FT METAL 101 101 IRON-SULFUR (2FE-2S).  
 FT METAL 104 104 IRON-SULFUR (2FE-2S).  
 FT METAL 208 208 IRON.  
 FT METAL 213 213 IRON.  
 FT METAL 362 362 IRON.  
 FT VARIANT 4 4 N->K (IN STRAIN G7).  
 FT VARIANT 12 12 S->F (IN STRAIN ATCC 17484).  
 FT VARIANT 15 15 S->T (IN STRAIN G7).  
 FT VARIANT 32 32 K->R (IN STRAIN G7).  
 FT VARIANT 50 50 A->S (IN STRAIN G7).  
 FT VARIANT 70 70 N->S (IN STRAIN G7).  
 FT VARIANT 90 91 SV->NA (IN STRAIN G7).

FT VARIANT 122 122 D->E (IN STRAIN G7).  
 FT VARIANT 173 173 M->I (IN STRAIN G7).  
 FT VARIANT 225 225 S->A (IN STRAIN G7).  
 FT VARIANT 225 225 S->C (IN STRAIN B8202).  
 FT VARIANT 232 232 A->V (IN STRAIN G7).  
 FT VARIANT 275 275 A->S (IN STRAIN G7).  
 FT VARIANT 391 391 E->R (IN STRAIN G7).  
 FT VARIANT 421 421 Q->K (IN STRAIN ATCC 17484).  
 FT VARIANT 434 434 H->D (IN STRAIN G7).  
 SQ SEQUENCE 449 AA; 49607 MW; 1ED2F4229684F7A8 CRC64;  
 Query Match 99.8%; Score 2404; DB 1; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 2,4e-181;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 MNYNKILVSSGSLQKRLIHGDELFQHEKTFEARNWLFTHDSLTPAGDYVAKMG 60  
 Db 1 MNYNKILVSSGSLQKRLIHGDELFQHEKTFEARNWLFTHDSLTPAGDYVAKMG 60  
 Oy 61 IDEVYSRQNDGSTRAPLNTGCRHKGKTLVSEAGNAGFCYSYHGMGFGSNGELQSVPE 120  
 Db 61 IDEVYSRQNDGSTRAPLNTGCRHKGKTLVSEAGNAGFCYSYHGMGFGSNGELQSVPE 120  
 Oy 121 KDLYGESLNKKKCLGLEYARVESFHGFTYGCQDQADPLMDYLGDAAMYLEPMFKHSGGL 180  
 Db 121 KDLYGESLNKKKCLGLEYARVESFHGFTYGCQDQADPLMDYLGDAAMYLEPMFKHSGGL 180  
 Oy 181 ELVGPBGVYVKAMWKPAEENFVDAYHVGWTHASSLSRSGESIFSSLAGNAALPREGAGL 240  
 Db 181 ELVGPBGVYVKAMWKPAEENFVDAYHVGWTHASSLSRSGESIFSSLAGNAALPREGAGL 240  
 Oy 241 QMTSKYSGGMVLMDGYSGVSHADLVPDLMAFGAKOERLNKEIGDVRARYSHLNTCY 300  
 Db 241 QMTSKYSGGMVLMDGYSGVSHADLVPDLMAFGAKOERLNKEIGDVRARYSHLNTCY 300  
 Oy 301 FPNNSMLTCSGVFYVWNPIDANTTEVWTYALVEKDPEDIKRLRLADSVQRTIPAGFWES 360  
 Db 301 FPNNSMLTCSGVFYVWNPIDANTTEVWTYALVEKDPEDIKRLRLADSVQRTIPAGFWES 360  
 Oy 361 DDNDMETSANGKRYGSRDLSNLGFGEDYGDVAIVGYGVKSAIGTYSRGFYRAY 420  
 Db 361 DDNDMETSANGKRYGSRDLSNLGFGEDYGDVAIVGYGVKSAIGTYSRGFYRAY 420  
 Oy 421 QAHVSSNMWAEFEHASSTWHTLTKTTDR 449  
 Db 421 QAHVSSNMWAEFEHASSTWHTLTKTTDR 449  
 RESULT 2  
 NDOB\_PSEFL STANDARD; PRT; 449 AA.  
 ID NDOB\_PSEFL  
 AC 007824;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (Naphthalene  
 1,2-dioxygenase ISP alpha).  
 GN NDOB OR NDOC2.  
 OS Pseudomonas fluorescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 17483;  
 RA Hamann C.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)  
 CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF  
 CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-  
 CC NAPHTHALENE DIHYDRODIOL.  
 CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) -> (1R,2S)-1,2-



CC dihydronaphthalene-1,2-diol + NAD(+).

CC -1- COPACITOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.

CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES

CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF

CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO

CC 1-HYDROXY-2-NAPHTHOL ACID AND THE METABOLISM OF DBT IS LIMITED TO

CC OXIDATION OF THE AROMATIC RING.

CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME

CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON

CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED

CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS

CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT

CC (NDOC).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

CC DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: AF004283; AAB61370.1; .

CC HSSP: P23094; INDO.

CC Interpro: IPR001281; Rieske.

CC Interpro: IPR001663; Ring\_hydroxyl\_A.

CC Pfam: PF00355; Rieske; 1.

CC Pfam: PF00848; Ring\_hydroxyl\_A; 1.

CC PRINTS: PR00090; RINGDIOXGNASE.

CC PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.

CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;

CC Dioxxygenase; NAD; plasmid.

CC METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 208 208 IRON (BY SIMILARITY).

CC METAL 213 213 IRON (BY SIMILARITY).

CC METAL 362 362 IRON (BY SIMILARITY).

CC SEQUENCE 449 AA; 49512 MW; 186A7924AD437A1C CRC64;

Query Match 96.3%; Score 2318; DB 1; Length 449;

Best Local Similarity 95.5%; Pred. No. 1.3e-174;

Matches 429; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 MNYNKKIIVSESGLSOKHLIHGDEELFOHELTIFARNKMLFTLHDSLIPAPGDVYATKMG 60

DB 1 MNYKKNKIIIVSESGLSOKHLIHGDEELFOHELTIFARNKMLFTLHDSLIPSPGDVYATKMG 60

QY 61 IDEVIVSRNDSDSIRAFNLVNCRRHCKTIVSVAGNAKGFVCSYHMGFGSNGELOSVPPE 120

DB 61 IDEVIVSRNDSDSIRAFNLVNCRRHCKTIVSVAGNAKGFVCSYHMGFGSNGELOSVPPE 120

QY 121 KDLYESLNNKKGGLGKEVARVESFHGFTIGCFDQAPPLMDYLGDAAWYLEMFKHSGGL 180

DB 121 KDLYESLNNKKGGLGKEVARVESFHGFTIGCFDQAPPLMDYLGDAAWYLEMFKHSGGL 180

QY 121 KDLYESLNNKKGGLGKEVARVESFHGFTIGCFDQAPPLMDYLGDAAWYLEMFKHSGGL 180

DB 121 KDLYESLNNKKGGLGKEVARVESFHGFTIGCFDQAPPLMDYLGDAAWYLEMFKHSGGL 180

QY 181 ELVGPGRKVIYANKKAPAEFNVGDAYHYGWTTHASSLSRGSIEFSSLAGNAALPREGAGL 240

DB 181 ELVGPGRKVIYANKKAPAEFNVGDAYHYGWTTHASSLSRGSIEFSSLAGNAALPREGAGL 240

QY 241 QMTSKYSGMGVLMGPGSVSHSADLVPELMARGAKQKQELNKETIDVYRIRSHLNCVY 300

DB 241 QMTSKYSGMGVLMGPGSVSHSADLVPELMARGAKQKQELNKETIDVYRIRSHLNCVY 300

QY 301 PPNNSMLTCSGVFKYWNPIDANTTEWYTAIYEKDMPEDLKRLADSVORTIGPAGFMES 360

DB 301 PPNNSMLTCSGVFKYWNPIDANTTEWYTAIYEKDMPEDLKRLADSVORTIGPAGFMES 360

QY 361 DDNDNMETASONGKYYQSHSDLSNLGEGEDVDYDAVYGVVGSAGIETSYRGFTYRAY 420

DB 361 DDNDNMETASONGKYYQSHSDLSNLGEGEDVDYDAVYGVVGSAGIETSYRGFTYRAY 420

DB 361 DDNDNMETASONGKYYQSHSDLSNLGEGEDVDYDAVYGVVGSAGIETSYRGFTYRAY 420

QY 421 QAHVSSSNMAEFHASTWHTLTKTTDR 449

DB 421 QAHVSSSNMAEFHASTWHTLTKTTDR 449

RESULT 3

NDOB\_PSEAE STANDARD; PRT; 449 AA.

AC 051494;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Naphthalene 1,2-dioxygenase alpha subunit (EC 1.14.12.12) (naphthalene

DE 1,2-dioxygenase ISP alpha).

GN NDOB OR PAHA3.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PAK1;

RA Takizawa N., Iida T., Yamauchi K., Satoh S., Wang Y., Fukuda M.,

RA Kiyohara H.;

RT "The molecular analysis of an NAH7-type gene cluster, pah, located on

RT the chromosome of Pseudomonas aeruginosa PAK1."

RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: COMPONENT OF NAPHTHALENE DIOXYGENASE (NDO)

CC MULTICOMPONENT ENZYME SYSTEM WHICH CATALYZES THE INCORPORATION OF

CC BOTH ATOMS OF MOLECULAR OXYGEN INTO NAPHTHALENE TO FORM CIS-

CC NAPHTHALENE DIHYDRODIOL.

CC -1- CATALYTIC ACTIVITY: Naphthalene + NADH + O(2) -> (1R,2S)-1,2-

CC dihydronaphthalene-1,2-diol + NAD(+).

CC -1- COPACITOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.

CC -1- PATHWAY: UPPER NAPHTHALENE CATABOLIC PATHWAY WHICH INVOLVES

CC CONVERSION OF NAPHTHALENE TO SALICYLATE, AND CATABOLISM OF

CC DIBENZOTHIOPHENE (DBT) AND PHENANTHRENE. CONVERTS PHENANTHRENE TO

CC 1-HYDROXY-2-NAPHTHOL ACID AND THE METABOLISM OF DBT IS LIMITED TO

CC OXIDATION OF THE AROMATIC RING.

CC -1- SUBUNIT: NAPHTHALENE DIOXYGENASE (NDO) MULTICOMPONENT ENZYME

CC SYSTEM IS COMPOSED OF AN ELECTRON TRANSFER COMPONENT AND AN IRON

CC SULFUR PROTEIN (ISP). THE ELECTRON TRANSFER COMPONENT IS COMPOSED

CC OF FERREDOXIN REDUCTASE (NDOR) AND FERREDOXIN (NDOA), AND ISP IS

CC COMPOSED OF A LARGE ALPHA SUBUNIT (NDOB) AND A SMALL BETA SUBUNIT

CC (NDOC).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING

CC DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: D84146; BAA12240.1; .

CC HSSP: P23094; INDO.

CC Interpro: IPR001281; Rieske.

CC Interpro: IPR001663; Ring\_hydroxyl\_A.

CC Pfam: PF00355; Rieske; 1.

CC Pfam: PF00848; Ring\_hydroxyl\_A; 1.

CC PRINTS: PR00090; RINGDIOXGNASE.

CC PROSITE: PS00570; RING\_HYDROXYL\_ALPHA; 1.

CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;

CC Dioxxygenase; NAD.

CC METAL 81 81 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 83 83 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 104 104 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

CC METAL 208 208 IRON (BY SIMILARITY).

FT METAL 213 213 IRON (BY SIMILARITY).  
 FT METAL 362 362 IRON (BY SIMILARITY).  
 SQ SEQUENCE 449 AA; 49715 MW; 35A189156722A21C CRC64;

Query Match 92.28; Score 2219; DB 1; Length 449;  
 Best Local Similarity 89.8%; Pred. No. 7.9e-167;

Matches 403; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

QY 1 MNNNKKLVSSSGLSOKHLIHGDEELFHELTETIFARWMLFTHDSLLIPAGDYATAMG 60  
 DB 1 MNNKMLVSSSGLTQKHLIHGDEELFHELTETIFARWMLFTHDSLLIPAGDYATAMG 60  
 QY 61 IDEVYVRONDGSIIRAFELNVCRIHKGKTLVSEAGNAKGFVSYHMGFGSGNELSVPEE 120  
 DB 61 VDEVYVRONDGSIIRAFELNVCRIHKGKTLVSEAGNAKGFVSYHMGFGSGNELSVPEE 120  
 QY 121 KDLYGSELNKKCLGKLEVARVESHGFTYGCFFDOBPAPLMYLGDAANYLEPMKHSGL 180  
 DB 121 KELYGELDKKCMKLEVARVESHGFTYGCFFDOBPAPLMYLGDAANYLEPMKHSGL 180  
 QY 181 ELVGPPEKVVYIKAMWKAPEAFVGVDAHVGMTTHASSLSRSGSIFSSLAGNAPLPEAGL 240  
 DB 181 ELVGPPEKVVYIKAMWKAPEAFVGVDAHVGMTTHASSLSRSGSIFSSLAGNAPLPEAGL 240  
 QY 241 QMTSKYSGMGVLMDGYSGVSHADLVPELMAFGAKOERLKEIGDYARITYRSHLNTCV 300  
 DB 241 QMTSKYSGMGVLMDGYSGVSHADLVPELMAFGAKOERLKEIGDYARITYRSHLNTCV 300  
 QY 301 FPNNSMLTSCGVFVWNPIDANTEVWTYAIVEKDMPEDLKRRLADSVORITGPAFWES 360  
 DB 301 FPNNSMLTSCGVFVWNPIDANTEVWTYAIVEKDMPEDLKRRLADSVORITGPAFWES 360  
 QY 361 DDNDMFTASONGKRYKSRDLSNLGFGEDYGDVAPGVGVSAGETSYRGEFYRAY 420  
 DB 361 DDNDMFTASONGKRYKSRDLSNLGFGEDYGDVAPGVGVSAGETSYRGEFYRAY 420  
 QY 421 QAHVSSNMAEFEEHASTWHTELTKTDR 449  
 DB 421 GAHSSSSMAEFEDVSKNWHTELAKTDR 449

RESULT 4  
 HCAE\_ECOLI STANDARD; PRT; 453 AA.  
 AC 047139; P77590; P78203;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 3-phenylpropionate dioxygenase alpha subunit (EC 1.14.1.-)  
 GN HCAE OR PHD1 OR HCAE OR HCA1 OR DIGA OR B2538 OR Z3809 OR ECS3404.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562, 83334;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RA Turlin E., Gasser F., Biville F.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12; MGI1655;  
 RX MEDLINE-97426617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN 13  
 RP SEQUENCE FROM N.A.

RC STRAIN-K12;  
 RX MEDLINE-97349980; PubMed-9205837;  
 RA Yamamoto Y., Alpha H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,  
 RA Oshima T., Oyama S., Saito N., Sakurai G., Satoh Y., Sivasubram S.,  
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata S., Horikuchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli  
 -K12 genome corresponding to 50.0-68.8 min on the linkage map and  
 RT analysis of its sequence features.";  
 RL DNA Res. 4:91-113(1997).  
 RN 14  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-0157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE-21074935; PubMed-11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grobeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN 15  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-0157:H7 / RIMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Ohnishi M., Murata T., Tanaka M., Tohe T.,  
 RA Kihara T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kudara S., Shiba T., Hattori M., Shigenaga H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- FUNCTION: PART OF MULTICOMPONENT 3-PHENYLPROPIONATE DIOXYGENASE.  
 CC CONVERTS 3-PHENYLPROPIONIC ACID (PP) INTO CIS-3-(3-CARBOXYETHYL)-  
 CC 3,5-CYCLOHEXADIENE-1,2-DIOL (PP-DIHYDRODIOL).  
 CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
 CC -1- PATHWAY: 3-phenylpropionate system catabolism.  
 CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE  
 CC FERREDOXIN (HCAC) AND A FERREDOXIN REDUCTASE (HCAD).  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
 CC DIOXYGENASE ALPHA SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/>  
 CC or send an email to [license@isb-stb.ch](mailto:license@isb-stb.ch)).  
 CC -----  
 DR EMBL; Z37966; CA86018.1; -;  
 DR EMBL; AE000340; AAC75591.1; -;  
 DR EMBL; D90883; BAA16433.1; -;  
 DR EMBL; D90884; BAA16441.1; -;  
 DR EMBL; AE005484; AAC57651.1; -;  
 DR EMBL; AP002562; BAB36827.1; -;  
 DR HSSP; P23094; INDO.  
 DR Ecogene; EG13456; hcae.  
 DR InterPro; IPR001281; Rieseke.  
 DR InterPro; IPR001663; Ring\_hydroxyl\_A.  
 DR Pfam; PF00355; Rieseke\_1.  
 DR Pfam; PF00848; Ring\_hydroxyl\_A; 1.  
 DR PRINTS; PR00090; RINGDIOXGNASE.  
 DR PROSITE; PS00570; RING\_HYDROXYL\_ALPHA; 1.  
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
 KW Dioxygenase; NAD; Complete proteome.  
 FT METAL 85 85 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 87 87 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 105 105 IRON-SULFUR (2FE-2S) (BY SIMILARITY).

```
FT METAL 108 108 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 213 213 IRON (BY SIMILARITY).
FT METAL 218 218 IRON (BY SIMILARITY).
FT CONFLICT 20 20 V -> A (IN REF. 1).
FT CONFLICT 364 453 GHRRNSKLCLEMGIGGOKRRDDGIPGITTNYIFSETARGM
FT YORADLLSESMEVDLKTAYAOEVMK -> ATAPATAN
FT CVMKGLVRSKSNATPAFLALLIISFOKIPLECTVAGDIF
FT (IN REF. 1).
SQ SEQUENCE 453 AA; 51109 MW; 02535BF5F463FD CMC64;

Query Match 30.6%; Score 737; DB 1; Length 453;
Best Local Similarity 36.4%; Pred. No. 1.9e-50;
Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGDELFOHEHKTIFARWMLFLTHDSLIPAPGDVYAKMGIDEVYSRONGSTRAPLN 79
DB 24 IYTDLDLQLELERFAWMLFLHDSQIPKPGDFNTYMGDAVAVVNAKDGSIKAFLN 83
QY 80 VCRHNGKTLVSEAGNAKGFVCSYHGMEFGSNGELQSVFEEDLYGESLNKKCLGKEVA 139
DB 84 QCRHRAMRVSADCGNTRAFTCPIYHOMWSYGINGLIDVLEFRATPQGLCKSHMKLINEVP 143
QY 140 RVSEHGFITGCFDOPAPPLMDYLDGDAWYLEPMF-KHSGLELVGPPGVVYKANMKAP 198
DB 144 CYESYKGLIFGMWDSAPGLRDYLDGDIAMYLDMDRREGTEIYGVOKVINCNMKFP 203
QY 199 AENFVGDANVGMTHASSL-----RSGESIFSLGNALP----- 234
DB 204 ABOFPASDOYHALFASHASAVQVLGANDGSKRLGD-----GQYARPVWETAKDALQFG 256
QY 235 --PEGAGLQMTSKYSGMGVLMDGYSGVHSAVLPELMAFGAKOERLKEIGDYARLY 292
DB 257 QGHSGGFEETFKPRANVAV--DGAVSSYRRETYAE-----AQGRIGEVRAIRL 303
QY 293 RSHLNTCFVPNNSMITCGSEVFYVWNPIDANTVEVYTAIVKQMPEDLKRRLADSVQRTI 352
DB 304 ACHNN--IFPLTSLWNGTFLTRVWHPRGDQVEVNAFCITTDKAASDEVAAEENSATRAF 361
QY 353 GRAGFESDDNMMETASNGKKYRSDSLNSLNGFSGEDVYGDVAVPGVYKSAIGETS 412
DB 362 GRAGFLEDSDSNKEIOKLLKGHRANSKLLCLEMGLGGERKRRDIPET--NYTFSETA 420
QY 413 YRGFRAYOAHVSSNMAEFEHASSTWHTELTK 445
DB 421 ARGWYGRMADLLSSSQEVLDTKTAAYOEVMK 453

RESULT 5
BED1_PSEPU STANDARD; PRT; 450 AA.
ID BED1_PSEPU 007944;
AC 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (EC 1.14.12.3).
GN BED01.
OS Pseudomonas putida.
OC Plasmid pMHT112.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN 11
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RC STRAIN-M12;
RX MEDLINE=93345820; PubMed=8344526;
RA Tan H.-M., Tang H.-Y., Joannou C., Abdel-Wahab N.H., Mason J.R.;
RT "The Pseudomonas putida M12 plasmid-encoded genes for benzene
dioxygenase are unusual in codon usage and low in G+C content.";
RL Gene 130:33-39(1993).
CC -1- CATALYTIC ACTIVITY: Benzene + NADH + O(2) -> cis-1,2-
dihydrobenzene-1,2-diol + NAD(+).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATECHOL.
```

```
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BEDC1 AND BEDC2), A
CC FERREDOXIN (BEDD) AND A FERREDOXIN REDUCTASE (BEDA).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF148496; AAL1758.1; -.
CC HSSP: P23094; INDO.
CC InterPro: IPR001281; Rieske.
CC InterPro: IPR001663; Ring_hydroxyl_A.
CC Pfam: PF00355; Rieske; 1.
CC Pfam: PF00848; Ring_hydroxyl_A; 1.
CC PRINTS: PR00090; RINGDIOXNAS.
CC PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
CC Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
CC Dioxxygenase; NAD; Plasmid.
CC METAL 96 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
CC METAL 222 222 IRON (BY SIMILARITY).
CC METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 51108 MW; 1ECDS5EA6C4C72C8 CMC64;

Query Match 27.1%; Score 652.5; DB 1; Length 450;
Best Local Similarity 34.6%; Pred. No. 8.3e-44;
Matches 148; Conservative 69; Mismatches 174; Indels 37; Gaps 11;

QY 20 IHGDELFOHEHKTIFARWMLFLTHDSLIPAPGDVYAKMGIDEVYSRONGSTRAPLN 79
DB 35 IYTDLDLQLELERFAWMLFLHDSQIPKPGDFNTYMGDAVAVVNAKDGSIKAFLN 94
QY 80 VCRHNGKTLVSEAGNAKGFVCSYHGMEFGSNGELQSVFEEDLYGESLNKKCLGKEVA 137
DB 95 QCRHNGKTLVSEAGNAKGFVCSYHGMEFGSNGELQSVFEEDLYGESLNKKCLGKEVA 148
QY 138 --VAVESFHGFITGCFDOPAPPLMDYLDGDAWYLEPMF-KHSGLELVGPPGVVYKANMKAP 194
DB 149 PLKARVETTYKGLIFANWDENAIDLPTYLGEAKFYMDHMDTEAGTEVYPIGLOKVVIPC 208
QY 195 WKAPENFVGDANVGMTHASSL-----RSGESIFSLGNALP----- 234
DB 209 WKFAEFCSDMYTAGTAAHLSGIIAGLPEDLELDLA--PP-----KFGQYRAS 257
QY 254 WDGY-SGVHSD-----LVPELMAF--GGAKOERLKEIGDYR--ARIYRSHLNTVF 301
DB 258 WGHSGSGFYIGDPNNMLMMGPKVSYLLEBPAEKAERIGSIERGKINLEHM--TVF 315
QY 302 PNNLSLTCGVFKVWNPIDANTVEVYTAIVKQMPEDLKRRLADSVQRTIGPAGFESD 361
DB 316 PTCSEFLPGVNTIRTWHPRGPEVWEVMAFTVVDADAPDDIKEEFRQTLRTFSAGVFEPD 375
QY 362 DNDNNETASQNGKKYOSRDSLNSLNGFSGEDVYGDVAVPGVYKSAIGETS 421
DB 376 DGENNVEIQHLIRGHKAKSRPPNMAEMSGQIVDNDPIYGRISNNVYSEEARGLYAHML 435
QY 422 AHVSSSNM 429
DB 436 KMMSPDW 443

RESULT 6
TOD1_PSEPU STANDARD; PRT; 450 AA.
ID TOD1_PSEPU P13450;
```

```

DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toluene 2,3-dioxygenase alpha subunit (EC 1.14.12.-).
GN TODC1.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-12.
RC STRAIN=F1;
RX MEDLINE=89359301; Pubmed=2670929;
RA Zylstra G.J., Gibson D.T.;
RT "Toluene degradation by Pseudomonas putida F1. Nucleotide sequence of
RT the todC12BAD genes and their expression in Escherichia coli.";
RL J. Biol. Chem. 264:14940-14946(1989).
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -1- PATHWAY: Toluene degradation; first step.
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (TODC1 AND TODC2), A
CC FERREDOXIN (TODB) AND A FERREDOXIN REDUCTASE (TODA).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: J04996; AAA26005.1; -
DR PIR: A36516; A36516.
DR HSSP: P23094; LINDO.
DR InterPro: IPR001281; Rleske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00335; Rleske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KW Dioxygenase; NAD.
FT METAL 96 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 IRON (BY SIMILARITY).
FT METAL 228 IRON (BY SIMILARITY).
SQ SEQUENCE 450 AA; 50944 MW; 038C80F197F3485D CRC64;
Query Match 27.0%; Score 650.5; DB 1; Length 450;
Best local Similarity 34.4%; Pred. No. 1.2e-43;
Matches 145; Conservative 71; Mismatches 101; Indels 25; Gaps 9;
QY 20 IHCDELFQHELKTIFARNMFLTHDSLIPAGDVTAKMIDEIVYSRQDGSIRAFLN 79
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 35 IYDEDLQLELERFARSMILLGHEIQIRKPGDYITTYMGEDPVVYVYRKDASIAVFLN 94
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 80 VCHRGKTVLVEGNKGFVCSFHGFGSNGELQSVPEFKDLGSLNKKCLGLAK- 137
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 95 QCHRRGRICRADGNKAFVCSYHGMAVYDAGNLVNPVYEAESFA-----CLNKKEWS 148
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 --VARVSEFGFTVGCDOEAPLMDYLGDAAWYLEPMF-KHSGLELVPGKVVITAN 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 149 PLKARVTVYGLLIFANDDENVDLDTYLGEAKFYMDHMLDPTETAGTEAIPGVQKVI PCN 208
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 195 WKAPAEVNDVAYHVGT-HASSLSRG--ESIFSLAGNALPPEGAGLQMTSKYSGMG 251
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 209 WKFAAEFCDDMYHAGTTSHTSLAGLAPEDL-----EMADLAPPTYGKQYRASMGIGS 263
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 252 VLMDGYSGVHSDLVPELMAF--GGAKQERLNKEIGDVR-ARIYRSHLNCTVPPNNSML 307
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 264 GFYVGDPMILMIAIMGPKVTSYWTGEPASERKAEBRLGSVERGSKIMVWHM--TVFPTCSFL 321
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 308 TCSGFKVWNPIDANTPEVWTVYAIVEKMDPEDLKRRLADSVORTISGAGWESDDNNME 367
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 322 PGINTRWHPRGPEVEVWMAFTVADADPDCLKEEFRROTLRFTFSAGVFEDDGNNVY 381
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 368 TASONGKRYQSRSDLSNLGFGEDVYGDAVPGVGSAGIGETSYRGFRAYQAHVSSS 427
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 382 EIGHLRGHKARSRPFNEMSMQTVNDPVPYGRISNNYSEBAANGLVAHMLRWMTSP 441
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 428 NW 429
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 442 DW 443
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 7
BPHA_BURCE STANDARD; PRT; 458 AA.
ID BPHA_BURCE
AC P37333;
DT 01-OCT-1994 (Rel. 30, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-
DE dioxygenase).
GN BPHA.
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LB400;
RX MEDLINE=92234948; Pubmed=1569021;
RX Erickson B.D., Mondello F.J.;
RA "Nucleotide sequencing and transcriptional mapping of the genes
RA encoding biphenyl dioxygenase, a multicomponent
RA polychlorinated-biphenyl-degrading enzyme in Pseudomonas strain
RA LB400.";
RL J. Bacteriol. 174:2903-2912(1992).
[2]
RN J. Bacteriol. 174:2903-2912(1992).
[3]
RP SEQUENCE OF 1-10, AND CHARACTERIZATION.
RC STRAIN=LB400;
RX MEDLINE=96011369; Pubmed=7592331;
RA Haddock J.D., Gibson D.T.;
RT "Purification and characterization of the oxygenase component of
RT biphenyl 2,3-dioxygenase from Pseudomonas sp. strain LB400.";
RL J. Bacteriol. 177:5834-5839(1995).
[3]
RN ERRATUM.
RA Haddock J.D., Gibson D.T.;
RP Haddock J.D., Gibson D.T.;
RL J. Bacteriol. 178:2158-2158(1996).
[1- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -1- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit.
CC -1- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -1- SUBUNIT: Heterohexamer consisting of three Bpha subunits and three
CC Bphe subunits. A ferredoxin (BphF) and a ferredoxin reductase
CC (BphC) must be present to obtain activity.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M86348; AAB63425.1; -
DR PIR: B41858; B41858.

```



```
Db      381 VPEODGEMWVLEQRYMRGRHKAKSTSLCAKMGILNPNKNKPVR---GATAYVYAEAR 437
Oy      415 GFYRAYQAHVSSSNW 429
Db      438 GMYHHWSRMMSSEPSW 452

RESULT 9
BPHA_PSEPS STANDARD: PRT: 458 AA.
ID BPHA_PSEPS Q52028;
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxygenase alpha subunit (EC 1.14.12.18) (Biphenyl 2,3-dioxygenase).
GN BPHA OR BPBH1.
OS Pseudomonas pseudoalcaligenes.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
OX NCBI_TaxID=330;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=KF707;
RX MEDLINE=92165849; PubMed=1537863;
RA Taira K., Hirose J., Hayashida S., Furukawa K.;
RT Analysis of bph operon from the polychlorinated biphenyl-degrading strain of Pseudomonas pseudoalcaligenes KF707";
RL J. Biol. Chem. 267:4844-4853(1992).
CC -I CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -I CORACTOR: Binds 1 2pe-2s cluster and 1 Iron atom per subunit (By similarity).
CC -I PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway; first step.
CC -I SUBUNIT: Heterohexamer consisting of three BphA subunits and three Bphe subunits. A ferredoxin (bphF) and a ferredoxin reductase (bphG) must be present to obtain activity (By similarity).
CC -I SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING DIOXYGENASE ALPHA SUBUNIT FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isdb-sib.ch/announce/or send an email to license@sib-sib.ch).
-----
DR EMBL, M83673; AAA25743.1; -.
DR HSSE, P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001633; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1..
DR Pfam: PF00848; Ring_hydroxyl_A_1.
DR PRINTS: PR00090; RNEDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA_1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron; Dioxygenase; NAD.
KW METAL 100
FT METAL 102 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 120 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 123 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 233 IRON (BY SIMILARITY).
FT METAL 239 IRON (BY SIMILARITY).
SO SEQUENCE 458 AA; 51437 MW; A6123AD09F97E482 CRC64;

Query Match 25.6%; Score 617; DB 1; Length 458;
Best Local Similarity 33.1%; Pred. NO. 5,2e-41;
Matches 147; Conservative 70; Mismatches 183; Indels 44; Gaps 13;
```

```

Db      31   EKGILDR-ITYADQSLYELELERYVGRSRLLLGHESHVEPTEDPLATYMGEDPVVMYRQK 89
Oy      71   DGSIRAEFLNVCNRHAKGTLYSVENAGNAKGVCSYHGMEFGSSNGELOSVPEPKDLYGESLNK 130
Dd      90   DKSIKVELNQCRRHGMRIQRSDAGNNKAFTCSYHGMAVDIAKLIVNVPREKEAFCDKREG 149
Oy      131  KQLGLKEV-----AVESFHEGFIYCFOFEAPRLMDYLGDAAWMLPEPFKRS-GGLELY 183
Dd      150  DC-GEPDKAEMGPLDAVAAYIKGLVFRANMVQARBDLETUIGDARPRMYDVLDRTPRAGTVAI 208
Oy      184  GPPGVKVIKANRKPAENFVGDAYHNG-VTHASSLSKSGEISFSLAGNALPREG--AGL 240
Dd      209  GDMOKKVVICCNKKFALEGSCDSMHNAGTMSHLISGLAG-----MPENDLSHA 256
Oy      241  QMTSKYSGSMGYLMDCYSVSHASADVPELMAFGAK-----QERLKEIGD 286
Dd      257  QVPTK-GNQFFRAGWGCHGSGWEDEFGMALVAMPKVTQYWTCEGPAAIDLAEORLGHTM-P 314
Oy      287  VARAIRSHLNCCTVPPNSMLTCSSGFVKWNPIDANTTEVVYAIVEKOMPEDLKRLAD 346
Dd      315  VR-RMGQM--SVPTCSFLPAINTIRFWHPRPETIEVWFMTLVDDADAPEIKEEYRR 371
Oy      347  SVQRTIGPAGFWSDDNDMMETASONGKKYGRSDLSNLSEFGEDVYGDADYPCVCGKS 406
Dd      372  NHIRTFSSAGVYEQDDGEWVELQKLRKYRAKKSQPLANQMKELGSRQSQTGRPFCNVG-Y 430
Oy      407  AIGETSYRGFYRAYQAHVSSSNWA 430
Dd      431  VYAEEARCGRMHMMRMSEPSPA 454

RESULT 10
BNZA_PSEPU ID BNZA_PSEPU STANDARD; PRI; 448 AA.
AC P08084;
Dt 01-AUG-1988 (Rel. 08, Created)
Dt 01-NOV-1988 (Rel. 09, Last sequence update)
Dt 15-JUL-1999 (Rel. 38, Last annotation update)
DE Benzene 1,2-dioxygenase alpha subunit (Ec 1.14.12.3) (P1 subunit).
GN BNZA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BE-81;
RX MEDLINE=88032840; PubMed=3667527;
RA Irie S., Doi S., Yorifuji T., Takagi M., Yano K.;
RT "Nucleotide sequencing and characterization of the genes encoding
benzene oxidation enzymes of Pseudomonas putida.";
RL J. Bacteriol. 169:5174-5179(1987).
CC -I- CATALYTIC ACTIVITY: Benzene + NAOH + O(2) = cis-1,2-
dihydrobenzene-1,2-diol + NAD(+).
CC -I- COFACTORS: PROBABLY BINDS A ZEP-ZS GROUP AND AN IRON ATOM.
CC -I- PATHWAY: FIRST STEP OF DEGRADATION OF BENZENE TO CATACHOL.
CC -I- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF FOUR PROTEINS: THE
TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BNZA AND BNZB), A
FERREDOXIN (BNZC) AND A FERREDOXIN REDUCTASE (BNZD).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation --
CC between the European Bioinformatics Institute. There are no restrictions on
its use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M17904; AAA25735.1; -.
DR PIR, A29830; A29830.
DR HSSP, P23094; INDO.
```

```
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxigenase; NAD.
FT METAL 96 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 98 98 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 116 116 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 119 119 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 222 222 IRON (BY SIMILARITY).
FT METAL 228 228 IRON (BY SIMILARITY).
SQ SEQUENCE 448 AA; 50208 MW; 6DDDBD5B2F040BE7 CRC64;

Query Match 24.4%; Score 588.5; DB 1; Length 448;
Best Local Similarity 33.6%; Pred. No. 8.8e-39;
Matches 142; Conservative 67; Mismatches 186; Indels 27; Gaps 11;

OY 20 IHGDELFQHELTIFARWMLFTLHDSLIAPAGDYTAAGIDEVYVSQNDGSIAPFLN 79
DB 35 IYDEDLVLELEFRVFAKSMLLGHETQIRKPEDYITTYMGEDPVVYVAKDASIAVFLN 94
OY 80 VCRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEFKDLGSELNKKCLGKEV- 137
DB 95 QCRHGRMCRISDADGNAKAFCTSYHGMAVDTAGNLVNVPEAESFA-----CLNKKEMS 148
OY 138 --VARSEHGFYGCFFDEAPPLMDYLGDAAWYLEPMF-KHSGGLELVGPPGKVIKAN 194
DB 149 PLKARETYKGLIFANMDENAVDLDTYLEAKYMDHMDRTEAGTEALPGVQKWVIPCN 208
OY 195 WKPAENFVDAYHVGWT-HASSLSRSG--ESTFSSLAGNAALPPBEAGIOMTSKXGSGM 251
DB 209 WKFAAQFQSDMTWAGTTHLSGILAGPEDL-----EMADLAPPPVGVQYRASKSGHGS 263
OY 252 VIMDGYSGVSHADLVPELMAF--GGAKEORLNKEIGDVR--ARIYSHLNCVTFEPNNSML 307
DB 264 GFVYGDPPNLMALIMGPXKVSYWTGEPASAKAERLGSVERGSLWMEHN--IYFPCFSL 321
OY 308 TCSGVKVNPIDANTTEVWTYAIYEKMPEDLKRLADSVORTIGPAGFWESDDDNNE 367
DB 322 PGINTVRLTASARERGEVWAFVVDADAPDDIKEEFRARL-RTEFSPA-CSSRTTGRTG 379
OY 368 TASONKKTYSRDSLSMLGFGEDYGVAVYPGVYKSAIGETSYRGYRYAVQAVSS 427
DB 380 SSSSTCEATSSRPNAKMSMDQYVNDPVTIGRISNNVYSBEARGLYAHMLRMVTS 439
OY 428 NW 429
DB 440 DW 441

RESULT 11
BPA1_PSES1 STANDARD; PRT; 458 AA.
ID BPA1_PSES1
AC 052438;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Biphenyl dioxigenase alpha subunit (BC 1.14.12.18) (biphenyl 2,3-
DE dioxigenase).
DE BPH1.
GN Pseudomonas sp. (strain KKS102).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=307;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=9432497; PubMed=8048958;
RA Fukuda M., Yasukouchi Y., Kikuchi Y., Nagata Y., Kimbara K.,
RA Horuchi H., Takagi M., Yano K.,
RT "Identification of the bphA and bphB genes of Pseudomonas sp. strains
RT KKS102 involved in degradation of biphenyl and polychlorinated
```

```
RT biphenyls.";
RL Biochem. Biophys. Res. Commun. 202:850-856(1994).
CC -I- CATALYTIC ACTIVITY: Biphenyl + NADH + O(2) = (2R,3S)-3-
CC phenylcyclohexa-3,5-diene-1,2-diol + NAD(+).
CC -I- COFACTOR: Binds 1 2Fe-2S cluster and 1 iron atom per subunit (By
CC similarity).
CC -I- PATHWAY: Biphenyl-polychlorinated biphenyl degradation pathway;
CC first step.
CC -I- SUBUNIT: Heterohexamer consisting of 3 BphA subunits and 3
CC BphA2 subunits. A ferredoxin (BphA3) and a ferredoxin reductase
CC (BphA4) must be present to obtain activity (By similarity).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: D17319; BAA04137.1; -.
DR HSP: P23094; INDO.
DR InterPro: IPR001281; Rieske.
DR InterPro: IPR001663; Ring_hydroxyl_A.
DR Pfam: PF00355; Rieske; 1.
DR Pfam: PF00848; Ring_hydroxyl_A; 1.
DR PRINTS: PR00090; RINGDIOXGNASE.
DR PROSITE: PS00570; RING_HYDROXYL_ALPHA; 1.
DR Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;
KM Dioxigenase; NAD.
FT METAL 101 101 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 103 103 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 121 121 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 124 124 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 234 234 IRON (BY SIMILARITY).
FT METAL 240 240 IRON (BY SIMILARITY).
SQ SEQUENCE 458 AA; 51876 MW; 128E8C38E2A3CFMA CRC64;

Query Match 24.4%; Score 587; DB 1; Length 458;
Best Local Similarity 33.6%; Pred. No. 1.2e-38;
Matches 147; Conservative 65; Mismatches 174; Indels 52; Gaps 14;

OY 20 IHGDELFQHELTIFARWMLFTLHDSLIAPAGDYTAAGIDEVYVSQNDGSIAPFLN 79
DB 40 IYADDDVLELEFRVFAKSMLLGHETQIRKPEDYITTYMGEDPVYVAKDASIAVFLN 99
OY 80 VCRHKGKTLVSEAGNAKGFVCSYHGMGFGSNGELQSVPEFKDLGSELNKKCLGKEV- 138
DB 100 QCRHGRMCRISDADGNAKAFCTSYHGMAVDTAGNLVNVPEAESFA-----CLNKKEMS 148
OY 139 -----ARVSEHGFYGCFFDEAPPLMDYLGDAAWYLEPMF-KHSGGLELVGPPGKVIK 192
DB 159 WGPLQAREVETKGLIFANMDAEPDLKTYLDAMPYMDMLDRTEAGTYVVGKQWVIP 218
OY 193 AMWKAPENFVDAYHVG-WTHASSLSRSGESTFSSLAGNAALPPBEAGIOMT-----SKG 247
DB 219 CWMKRAEQFQSDMTWAGTTHLSGILAGPEDL-----EMADLAPPPVGVQYRASKSGHGS 263
OY 248 SGMGYLMDGYSGVSHADLVPELMAFGAK-----QERLNKEIGDVRAR-IYRSHL 296
DB 264 SGPRAAGHSGCWFTINDAAILMAVGRITTYWOGPAAEAKRKLNDMPOTMFGQIM 323
OY 297 NCTVFPNNSMLTCSGVKFWNPIDANTTEV---WTYAIYEKMPEDLKRLADSVORTIG 353
DB 324 --TVFPCTGFLPGINTIRSWHPRGNEVECGPSWMSMPAR---PEDIKEFRONIRTFN 378
OY 354 PGFWESDDNDNMFAASQKKTYSRDSLSMLGFGEDYGVAVYPGVYKSAIGETSYRGYRY 411
DB 379 AGTTEDDGDEWWEIQRGLRGHAKSAPLCAOMGLNVNPKSNPDF---GKTAVYAE 435
OY 412 SYRGFYRAVQAVHSSNM 429
```

Db 436 AARGMYHHWARMSEPNW 453

## RESULT 12

XYLX\_PSEPU  
ID XYLX\_PSEPU STANDARD; PRT; 454 AA.  
AC P23099;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Toluate 1,2-dioxygenase alpha subunit (EC 1.14.12.-).  
GN XYLX.  
OS Pseudomonas putida.  
OC Plasmid TOL plasmid.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas  
OX NCBI\_taxid=303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92155191; Pubmed=1740120;  
RA Neidle E.L., Hartnett C., Ornstom L.N., Bairoch A., Reikik M.,  
RA Harayama S.,  
RT C15-diol dehydrogenases encoded by the TOL plasmid xylX gene  
RT and the Acinetobacter calcoaceticus chromosomal bend gene are members  
RT of the short-chain alcohol dehydrogenase superfamily."  
RL Eur. J. Biochem. 204:113-120(1992).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS:  
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (XYLX AND XYLX'), AND AN  
CC ELECTRON TRANSFER COMPONENT (XYLZ).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M64747; AAA26047.1; -  
DR PIR; S23482; S23482.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00848; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS; PS00570; RING-HYDROXYL.ALPHA; 1.  
DR PROSITE; PS00570; RING-HYDROXYL.ALPHA; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;  
KM Dioxigenase; NAD; Plasmid.  
FT METAL 92 92 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 94 94 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
FT METAL 221 221 IRON (BY SIMILARITY).  
FT METAL 226 226 IRON (BY SIMILARITY).  
SQ SEQUENCE 454 AA; 51898 MW; F68DC41E236C077 CRC64;

Query Match 17.0%; Score 410.5; DB 1; Length 454;  
Best Local Similarity 31.2%; Pred. No. 8.5e-25;  
Matches 105; Conservative 64; Mismatches 135; Indels 33; Gaps 14;

QY 23 DEELFOHELTETEARNWLETHSLIPAPGDYVAKMGIDIVISRONDSIRAFANVC 82  
DB 34 DPLFLEEMHIIFEGNNIYLAEHSQIPEKNDYTTTOMGRQPIFTRNKGELAEVFNAC 93  
QY 83 HRCGLTVSAGNAKGVCSHGNGSNGELQVPRPKD-LYGESLNKKCLG---LKEV 138  
DB 94 HRGATLCFRSGNAKATHTCSHGNTFSNGSKLLAVKDPKAGAGYDPSFD--CDGSHDLKRV 151

QY 139 ARVESFHGFIYGCPEQDEAPPLMDYLGDANWLEPFMRHS--GGELEVGPPKRVYIKANMKA 197  
DB 152 ARFASVGRGFLGSLREDVAPLEEFEGSRKVIDWVVOGSGLELVLAGSSYYVEGMKV 211  
QY 198 PAENFVGAIVHG--WTHASS-----LR-SGESIFSSLAGNAALPPEGAGLQMTSKYGS 248  
DB 212 QVEN-GADGYHVSFTVHWNYAATQOQRKLRDAGDIDRAMTA--SSWGDGGGCFY---SFEN 265  
QY 249 GNGVIMDQSGVSHADVPELMAFGAKOEKRLNKGIDVRR-LYRSHLQCTVEPNNSML 307  
DB 266 GHOWMARMGDKPNPLP-----AERDLASEFGEARDMIGVSRNLCLPNTLYLM 317  
QY 308 TCSG-VFKWNPIDANTTEWTVTAIVEK-DMPEDLKR 342  
DB 318 DQFGSLRITRPLSVDRPRTITTYCIAPRGEPFRARR 354

## RESULT 13

BENA\_ACICA  
ID BENA\_ACICA STANDARD; PRT; 461 AA.  
AC P07769;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10).  
GN BENA.  
OS Acinetobacter calcoaceticus.  
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
OC Acinetobacter.  
OX NCBI\_taxid=471;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BD413 / ADP1;  
RX MEDLINE=91358314; Pubmed=1885518;  
RA Neidle E.L., Hartnett C., Ornstom L.N., Bairoch A., Reikik M.,  
RA Harayama S.;  
RT "Nucleotide sequences of the Acinetobacter calcoaceticus benABC genes  
RT for benzoate 1,2-dioxygenase reveal evolutionary relationships among  
RT multicomponent oxygenases."  
RL J. Bacteriol. 173:5385-5395(1991).  
RN [2]  
RP REVISIONS TO 84; 103-104; 171-172 AND 380-382.  
RP Elby D.M., Neidle E.L.;  
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: DEGRADATION OF BENZOATE TO 2-HYDRO-1,2-DIHYDROXYBENZOATE  
CC (DHB).  
CC -1- CATALYTIC ACTIVITY: Benzoate + NADH + O(2) = catechol + CO(2) +  
CC NAD(+).  
CC -1- COFACTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.  
CC -1- PATHWAY: DEGRADATION PATHWAY LEADING FROM BENZOATE TO CATECHOL.  
CC -1- SUBUNIT: THIS DIOXYGENASE SYSTEM CONSISTS OF THREE PROTEINS: THE  
CC TWO SUBUNITS OF THE HYDROXYLASE COMPONENT (BENA AND BEMB), AND  
CC AN ELECTRON TRANSFER COMPONENT (BENC).  
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING  
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF009224; AAC46436.2; -  
DR PIR; S23477; S23477.  
DR InterPro: IPR001281; Rieske.  
DR InterPro: IPR001663; Ring\_hydroxyl\_A.  
DR Pfam: PF00848; Rieske; 1.  
DR Pfam: PF00848; Ring\_hydroxyl\_A; 1.  
DR PRINTS; PS00570; RING-HYDROXYL.ALPHA; 1.  
DR PROSITE; PS00570; RING-HYDROXYL.ALPHA; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; Iron-sulfur; Iron;



```

KW Dioxigenase; NAD.
FT METAL 95 95 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 97 97 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 115 115 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 118 118 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 224 224 IRON (BY SIMILARITY).
FT METAL 229 229 IRON (BY SIMILARITY).
SQ SEQUENCE 461 AA; 52228 MW; 3CFC3247A3C4379 CRC64;

Query Match 16.6%; Score 399; DB 1; Length 461;
Best Local Similarity 27.7%; Pred. No. 7e-24;
Matches 128; Conservative 78; Mismatches 204; Indels 52; Gaps 20;

OY 5 NRIIV--SESGLSOKH-LIHGDELFQHEHLEKTFARMMLETHSLIPAGDYVYAKMG 61
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db DELVDNMTGEGFKLRSVFTDQALFDLEMKYIFEGSNWVYLAHESQIPNNNYTTIYICR 75

OY 62 DEVIYSRQNDGSIKRAFLNVCRRHKTLYSVEGNAKGEVCSYHMGFGSNGELQSVPEK 121
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db QPILARNNGELNMINACSHRGAQLCHRKGRNKTYYTCPPHGWTFNNSGKLLKYKDP 135

OY 122 DL-YGSLUNK-CLGKEXARVESFHGFYGCQDEAPRLMYLGDAAVLEPMKHS- 178
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 136 DNGYSDCFQDQSDSHDKVARESTKGFGLGSLNPDVPSLQFLBETTKIIMIVGQSDQ 195

OY 179 GELVGPPEKVVYIKANKAPAEFVGDATYHG--WTHASL-----RSGSIFSSLAG 229
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 196 GLEVLGVSTYTYEGNMKLTAE--GADGYHSAVHNMYATQHKREKAGOTIRAMSG 254

OY 230 NAALPEGAGLOMTSKYG--SGMGVLMDSYGSVHSDVPELM---AFGAKOEHLNKE 283
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 255 SWG--KHGGG-----SYGEHGHMLMTQWGNEDRPNPKAEYTEKGAAMSKM--- 304

OY 284 IGDVARIYRSHLNCVPEPNNMLTCSG--VFKVWNPIDANTPEVMTYALVE-KDMPEDLK 341
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 305 -----IERSR-NICLFPNYLMDQFGSQIVLRISYVKTEVITYICAPVGEAPEARA 356

OY 342 RLADSVORTTIPAGFWESDDNDNMETASQNGKRYTSRDSRLSLNFGEDVYG--DA-- 397
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 357 RLIR-QYEDFFNAGSMATPDDLEFPACQAGYAGIELENDMCR--GSKMHWIGPDDAAN 413

OY 398 ---VYPGVVGSKAIGETSTRGYRAYQAHVSSNMAPEFHAS 436
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 414 EIGLKPALSGIKTEDEGLYLAOHYWLKSMQOIAAEKEEFAS 455

RESULT 14
YEAM_ECOLI
ID YEAM_ECOLI STANDARD: PRT: 374 AA.
AC P76253;
DR 15-JUL-1998 (Rel. 36, Created)
DR 15-JUL-1998 (Rel. 36, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative dioxigenase alpha subunit yeam (EC 1.14.1.-).
GN YEAM OR B1802 OR Z2845 OR ECS2511.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
NCBI_TaxID=562, 83334;
RN NCBI_TaxID=562, 83334;
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;

```

```

RX MEDLINE=97251358; PubMed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizubuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,
RA Sivasubdam S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horiiuchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40.1-50.0 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller U.,
RA Grodbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).

RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohsudo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubota S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).

CC -I- CORAFCTOR: PROBABLY BINDS A 2FE-2S GROUP AND AN IRON ATOM.
CC -I- SUBUNIT: PROBABLE HETERODIMER OF YEAM AND YEAX.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL RING-HYDROXYLATING
CC DIOXYGENASE ALPHA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AE000274; AAC74872.1; -
DR EMBL, D90823; BAA15597.1; ALT_INIT.
DR EMBL, D90824; BAA15606.1; ALT_INIT.
DR EMBL, AE005403; AAG56791.1; -
DR EMBL, AP002558; BAB35934.1; -
DR Ecocore, EG13509; yeam.
DR InterPro, IPR001281; Rieske.
DR InterPro, IPR001663; Ring_hydroxyl_A.
DR Pfam, PF00355; Rieske; 1.
DR PROSITE, PS00570; RING_HYDROXYL_ALPHA; FALSE NEG.
KW Hypothetical protein: Oxidoreductase; Iron-sulfur; Iron; Dioxigenase;
KW NAD; Complete proteome.
FT METAL 89 89 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 91 91 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 109 109 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 112 112 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 211 211 IRON (BY SIMILARITY).
FT METAL 216 216 IRON (BY SIMILARITY).
SQ SEQUENCE 374 AA; 42561 MW; BB5386ACA9585606 CRC64;

Query Match 10.3%; Score 249; DB 1; Length 374;
Best Local Similarity 26.6%; Pred. No. 3.2e-12;
Matches 98; Conservative 54; Mismatches 126; Indels 90; Gaps 19;

OY 23 DEELFQHEKLTIFARNWLEFLTHDSLIPAGDYVYAKMGIDEVYSRQNDGSIKRAFLNVCRR 82
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 31 DONAFHEKENVFASKSWICVASHSELNANDVYVTRIIIGESIVLVGRDKVLRAPYVNC 90

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:55:37 ; Search time 13.1667 seconds  
(without alignments)  
3278.305 Million cell updates/sec

Title: US-09-843-250-36

Perfect score: 2408

Sequence: 1 MNYNKILVSESGLSQKHLI.....AEFFHASSTWHTLTKTDR 449

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2404	99.8	449	2 JS0071	naphthalene dioxyg
2	2404	99.8	462	2 S27632	naphthalene 1,2-di
3	2346	97.4	449	2 JN0644	naphthalene 1,2-di
4	2292	95.2	449	2 C55217	polycyclic aromati
5	2056	85.4	447	2 JC5352	2-nitrotoluene dio
6	1957	39.7	459	2 T31134	naphthalene dioxyg
7	774	32.1	450	2 T31256	terminal oxygenase
8	737	30.6	453	2 A65031	biphenyl dioxygena
9	737	30.6	453	2 D91054	biphenyl dioxygena
10	725.5	30.1	455	2 G85898	biphenyl dioxygena
11	725.5	27.1	450	2 T31258	aromatic oxygenase
12	652.5	27.0	450	1 JN0812	benzene 1,2-dioxyg
13	650.5	26.5	461	2 A36516	toluene dioxygenas
14	639	26.5	461	2 S51757	biphenyl dioxygena
15	617.5	22.6	459	1 B41858	biphenyl dioxygena
16	617	22.6	457	1 JC4993	biphenyl dioxygena
17	617	22.6	458	1 A42409	biphenyl dioxygena
18	605	25.1	431	2 JN0098	carbazole dioxygen
19	588.5	24.4	448	1 A29830	benzene 1,2-dioxyg
20	587	24.4	458	2 JC3467	biphenyl dioxygena
21	452	18.8	469	2 T50934	dioxygenase dihal,
22	416	17.3	471	2 T31281	benzoate 1,2-dioxy
23	410.5	17.0	454	2 A41659	benzoate 1,2-dioxy
24	409.5	17.0	455	2 E83332	toluene 1,2-dioxyg
25	394	16.4	461	2 S23477	probable benzoate
26	391	16.2	464	2 G83331	anthranilate dioxy
27	341.5	14.2	424	2 E83384	probable ring-hydr
28	333	13.8	426	2 T31278	biphenyl dioxygena
29	315.5	13.1	391	2 T31251	aromatic oxygenase

30	311.5	12.9	468	2 G97447	hypothetical prote
31	311.5	12.9	468	2 A82665	ring hydroxylating
32	296	12.3	429	2 F82970	probable ring hydr
33	287	11.9	412	2 H95311	probable aromatic-
34	283	11.8	415	2 A63320	benzoate 1,2-dioxy
35	270.5	11.2	404	2 H87635	Rieske 2Fe-2S faml
36	268	11.1	420	2 T31285	biphenyl dioxygena
37	252.5	10.5	374	2 AF0304	probable dioxygena
38	249	10.3	374	2 C85791	probable choline m
39	249	10.3	374	2 G90942	probable choline m
40	249	10.3	374	2 B64941	probable choline m
41	233.5	9.7	382	2 G70946	probable dioxygena
42	224	9.3	446	2 T14542	choline monooxygen
43	223	9.3	439	2 T09214	choline monooxygen
44	214.5	8.9	186	2 JC5354	2-nitrotoluene dio
45	186	7.7	426	2 T08550	choline monooxygen

## ALIGNMENTS

```
RESULT 1
JS0071
naphthalene dioxygenase (EC 1.14.12.-) ndoB protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 20-Aug-1999
C:Accession: JS0071
R:Kurkela, S.; Leivaeslahti, H.; Palva, E.T.; Teerl, T.H.
Gene 73, 355-362, 1988
A:Title: Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase system is composed of three proteins.
A:Reference number: JS0070; MUID:89211973; PMID:3243438
A:Accession: JS0071
A:Molecule type: DNA
A:Residues: 1-449 <KUR>
A:Cross-References: GB:M23914; NID:g151392; PID:NAB47591.1; PID:g151394
C:Comment: Naphthalene dioxygenase system is composed of three proteins.
C:Genetics:
A:Gene: ndoB
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske (
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein
F:71-119/Domain: Rieske (2Fe-2S) homology <RSK>
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status
```

Query Match 99.8% Score 2404; DB 2; Length 449;  
Best Local Similarity 99.8%; Pred. No. 4e-184;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MNYNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLIPAPGDYATKMG	60
DB	1	MNYNKILVSESGLSQKHLIHGDEELFOHELKTIFARNMLFTLHDSLIPAPGDYATKMG	60
QY	61	IDEVIVSRONDGSIKAFNLVNCRRGKTLVSVBGNKGFVCSYHGMFGSGNELSVPRE	120
DB	61	IDEVIVSRONDGSIKAFNLVNCRRGKTLVSVBGNKGFVCSYHGMFGSGNELSVPRE	120
QY	121	KLTGSESLNKKCLGKEVAVRVSFHEITGCFDQEPRLMDVLGDAAWLEPMFRHSGGL	180
DB	121	KLTGSESLNKKCLGKEVAVRVSFHEITGCFDQEPRLMDVLGDAAWLEPMFRHSGGL	180
QY	181	ELVGPGRKVIKANKKAPAFENFVGDAVHYGWTNASSLSRSGESIFSSLAGNALPPEGAGL	240
DB	181	ELVGPGRKVIKANKKAPAFENFVGDAVHYGWTNASSLSRSGESIFSSLAGNALPPEGAGL	240
QY	241	QMTSKYSGMGVLMQGYSCVHSADLVPELMATGAKOEKRLNKEIGDVARIRYSHLNCYV	300
DB	241	QMTSKYSGMGVLMQGYSCVHSADLVPELMATGAKOEKRLNKEIGDVARIRYSHLNCYV	300
QY	301	PNNSMLTCSGVFKYWNPLDANTTEWYTAIVEKMPEDLKRRLADSVORTGPGAFWES	360
DB	301	PNNSMLTCSGVFKYWNPLDANTTEWYTAIVEKMPEDLKRRLADSVORTGPGAFWES	360
QY	361	DDNDNMTASQNGKRYOSRDSGLSNLGFGEVDYGDVAVPGVYGSALGETSYRGFYRAY	420
DB	361	DDNDNMTASQNGKRYOSRDSGLSNLGFGEVDYGDVAVPGVYGSALGETSYRGFYRAY	420

Db 361 DDNDNMTASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420

OY 421 QAHVSSNMMAEFHASSSTWHTLTKTTDR 449

Db 421 QAHVSSNMMAEFHASSSTWHTLTKTTDR 449

## RESULT 2

S27632  
naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large chain  
C:Species: Pseudomonas sp.  
C:Date: 06-Jan-1995 #sequence-revision 06-Jan-1995 #text-change 15-Oct-1999  
C:Accession: S27632; B49343  
R:Denome, S.A.; Young, K.D.  
Submitted to the EMBL Data Library, February 1992  
A:Description: Cloning and molecular characterization of genes involved in metabolism of  
A:Reference number: S27631  
A:Accession: S27632  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <DEN>  
A:Cross-references: EMBL:M60405  
R:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
J. Bacteriol. 175, 6890-6901, 1993  
A:Title: Metabolism of dibenzothiophene and naphthalene in Pseudomonas strains: complete  
A:Reference number: A49343; MUID:94042852; PMID:8226631  
A:Accession: B49343  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 14-462 <DE2>  
A:Cross-references: GB:M60405; NID:g151195; PIDN:AAA16125.1; PID:g294351  
A:Experimental source: strain C18  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:84-112/Domain: Rieske [2Fe-2S] homology <RSK>  
F:94,96,114,117/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 99.8%; Score 2404; DB 2; Length 462;  
Best Local Similarity 99.8%; Pred. No. 4,2e-184;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MNYNKKILVSSGSLQKHLHGDEELFQHEKLTFFARNWMLFTHDSLIPAPGDVYAKMG 60  
Db 14 MNYNKKILVSSGSLQKHLHGDEELFQHEKLTFFARNWMLFTHDSLIPAPGDVYAKMG 73

OY 61 IDEVYSRONGSGIRAFPLNVCRRHGKTLVSEAGNAKGFVCSYHGWFSGNGELOSYPFE 120  
Db 74 IDEVYSRONGSGIRAFPLNVCRRHGKTLVSEAGNAKGFVCSYHGWFSGNGELOSYPFE 133

OY 121 KDLYGESLNKCKLGLKEVARVSEFHGFIYGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL 180  
Db 134 KDLYGESLNKCKLGLKEVARVSEFHGFIYGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL 193

OY 181 ELVGPCKVYIKAMWKAPAEVNDAYHVGWTHASSLSRSGESTFSSLAGNAALPPREGAGL 240  
Db 194 ELVGPCKVYIKAMWKAPAEVNDAYHVGWTHASSLSRSGESTFSSLAGNAALPPREGAGL 253

OY 241 QMTSKYSGMGVLMDGSGVHSADLVPDLMAFGAKOERLKEIGDVARITYRSHLNCYV 300  
Db 254 QMTSKYSGMGVLMDGSGVHSADLVPDLMAFGAKOERLKEIGDVARITYRSHLNCYV 313

OY 301 FPNNSMILTCGSGVFKVWNPIDANTTEVMTYALVEKDMPEDLKRRRLADSVQRTIGPAGFWES 360  
Db 314 FPNNSMILTCGSGVFKVWNPIDANTTEVMTYALVEKDMPEDLKRRRLADSVQRTIGPAGFWES 373

OY 361 DDNDNMTASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 374 DDNDNMTASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 433

OY 421 QAHVSSNMMAEFHASSSTWHTLTKTTDR 449  
Db 434 QAHVSSNMMAEFHASSSTWHTLTKTTDR 462

## RESULT 3

JN0644  
naphthalene 1,2-dioxygenase (EC 1.14.12.12) iron-sulfur oxygenase component large cha  
N:Alternate names: nahc protein  
C:Species: Pseudomonas putida  
C:Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text-change 20-Aug-1999  
C:Accession: JN0644  
R:Simon, M.J.; Ossilund, T.D.; Saunders, R.; Ensley, B.D.; Suggs, S.; Harcourt, A.; Su  
Gene 127, 31-37, 1993

A:Title: Sequences of genes encoding naphthalene dioxygenase in Pseudomonas putida st  
A:Reference number: JN0640; MUID:93252277; PMID:8486285  
A:Accession: JN0644  
A:Molecule type: DNA  
A:Residues: 1-449 <STM>  
A:Cross-references: GB:M83949; NID:g151384; PIDN:AAA25902.1; PID:g151387  
C:Comment: This protein is a member of naphthalene dioxygenase multicomponent enzyme  
C:Genetics:  
A:Gene: nahc  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2  
C:Keywords: 2Fe-2S; heterotrimer; metalloprotein; oxidoreductase; Rieske iron-sulfu  
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 97.4%; Score 2346; DB 2; Length 449;  
Best Local Similarity 96.7%; Pred. No. 1,7e-179;  
Matches 434; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

OY 1 MNYNKKILVSSGSLQKHLHGDEELFQHEKLTFFARNWMLFTHDSLIPAPGDVYAKMG 60  
Db 1 MNYNKKILVSSGSLQKHLHGDEELFQHEKLTFFARNWMLFTHDSLIPAPGDVYAKMG 60

OY 61 IDEVYSRONGSGIRAFPLNVCRRHGKTLVSEAGNAKGFVCSYHGWFSGNGELOSYPFE 120  
Db 61 IDEVYSRONGSGIRAFPLNVCRRHGKTLVSEAGNAKGFVCSYHGWFSGNGELOSYPFE 120

OY 121 KDLYGESLNKCKLGLKEVARVSEFHGFIYGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL 180  
Db 121 KELYGESLNKCKLGLKEVARVSEFHGFIYGCFODEAPPLMDYLGDAAMYLEPMFKHSGGL 180

OY 181 ELVGPCKVYIKAMWKAPAEVNDAYHVGWTHASSLSRSGESTFSSLAGNAALPPREGAGL 240  
Db 181 ELVGPCKVYIKAMWKAPAEVNDAYHVGWTHASSLSRSGESTFSSLAGNAALPPREGAGL 240

OY 241 QMTSKYSGMGVLMDGSGVHSADLVPDLMAFGAKOERLKEIGDVARITYRSHLNCYV 300  
Db 241 QMTSKYSGMGVLMDGSGVHSADLVPDLMAFGAKOERLKEIGDVARITYRSHLNCYV 300

OY 301 FPNNSMILTCGSGVFKVWNPIDANTTEVMTYALVEKDMPEDLKRRRLADSVQRTIGPAGFWES 360  
Db 301 FPNNSMILTCGSGVFKVWNPIDANTTEVMTYALVEKDMPEDLKRRRLADSVQRTIGPAGFWES 360

OY 361 DDNDNMTASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420  
Db 361 DDNDNMTASONGKRYOSRSDLSNLGFGEDVYGDVAYPGVYKSAIGETSYRGFYRAY 420

OY 421 QAHVSSNMMAEFHASSSTWHTLTKTTDR 449  
Db 421 QAHVSSNMMAEFHASSSTWHTLTKTTDR 449

## RESULT 4

C55217  
polycyclic aromatic hydrocarbon dioxygenase (EC 1.14.12.-) iron-sulfur protein large  
C:Species: Pseudomonas putida  
C:Date: 05-May-1995 #sequence-revision 05-May-1995 #text-change 20-Jun-2000  
C:Accession: C55217  
R:Takizawa, N.; Kaida, N.; Toriogo, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara  
J. Bacteriol. 176, 2444-2449, 1994  
A:Title: Identification and characterization of genes encoding polycyclic aromatic hy  
A:Reference number: A55217; MUID:94209249; PMID:8157615  
A:Accession: C55217  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-449 <TAG>  
A:Cross-references: GB:AB004059; GB:D16629; NID:92189972; PIDN:BA020391.1; PID:9391844  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:71-119/Domain: Rieske [2Fe-2S] homology <RSK>  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 95.2%; Score 2292; DB 2; Length 449;  
Best Local Similarity 94.4%; Pred. No. 3.5e-175;  
Matches 424; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNYNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWLEFLTHDSLIPAPGDVYTAAMG 60  
DB 1 MNYNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWLEFLTHDSLIPAPGDVYTAAMG 60  
QY 61 IDEVIVSRONQSGIRAFINVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELOSVPFE 120  
DB 61 IDEVIVSRONQSGIRAFINVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELOSVPFE 120  
QY 121 KDLYGESLNKKKCLGKEVARVESFHGFYTCGFDDEAPPLMDYLDGDAAMYLEPMPKHSGL 180  
DB 121 KELVYESLNKKKCLGKEVARVESFHGFYTCGFDDEAPPLMDYLDGDAAMYLEPMPKHSGL 180  
QY 181 ELVGPFGKVVYIKAMKKAPENFVGDYHVGTWTHASLSRSGESIFSSLAGNALPPEGAGL 240  
DB 181 ELVGPFGKVVYIKAMKKAPENFVGDYHVGTWTHASLSRSGESIFSSLAGNALPPEGAGL 240  
QY 241 QMTSKYSGMGVLMDGYSVSHADLVPELMAFGAKQERLKEIGDVARIYRSHLNTCV 300  
DB 241 QMTSKYSGMGVLMDGYSVSHADLVPELMAFGAKQERLKEIGDVARIYRSHLNTCV 300  
QY 301 FPNNSMLTCSGVFWVWNPIDANTTEVMTYATVEKMPEDLKRRLADSVQRTIGPAGFMS 360  
DB 301 FPNNSMLTCSGVFWVWNPIDANTTEVMTYATVEKMPEDLKRRLADSVQRTIGPAGFMS 360  
QY 361 DDNDNMTASONGKRYOSRDSLSNLGFGEDVYGDVAVPGVKSALIGETSYRGFYRAY 420  
DB 361 DDNDNMTASONGKRYOSRDSLSNLGFGEDVYGDVAVPGVKSALIGETSYRGFYRAY 420  
QY 421 QAHVSSNMWAEFEHASTWHTLTKTTDR 449  
DB 421 QAHVSSNMWAEFEHASTWHTLTKTTDR 449

## RESULT 5

JC5352  
2-nitrotoluene dioxygenase (EC 1.14.12.-) iron-sulfur protein large chain - Pseudomonas  
N:Alternate names: 2-nitrotoluene iron-sulfur protein alpha chain  
C:Species: Pseudomonas sp.  
C:Date: 27-May-1997 #sequence\_revision 18-Jul-1997 #text\_change 20-Aug-1999  
R:Accession: JC5352  
R:Paralogs: J.V.; Kumar, A.; Parales, R.E.; Gibson, D.T.  
Gene 181, 57-61, 1996  
A:Title: Cloning and sequencing of the genes encoding 2-nitrotoluene dioxygenase from Ps  
A:Reference number: JC5350; MUID:97128768; PMID:89733308  
A:Accession: JC5352  
A:Molecule type: DNA  
A:Residues: 1-447 <PAR>  
A:Cross-references: GB:049504; NID:91773273; PIDN:AA040383.1; PID:91773277  
A:Experimental source: strain JS42  
C:Comment: This enzyme catalyzes the addition of dioxygen to the aromatic nucleus of 2-n  
C:Genetics:  
A:Gene: ntdAc  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:69-117/Domain: Rieske [2Fe-2S] homology <RSK>  
F:79,81,99,102/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 85.4%; Score 2056; DB 2; Length 447;  
Best Local Similarity 84.0%; Pred. No. 2.5e-156;  
Matches 377; Conservative 34; Mismatches 36; Indels 2; Gaps 1;

QY 1 MNYNKILVSESGLSQKHLIHGDEELFQHELKTIFARNWLEFLTHDSLIPAPGDVYTAAMG 60  
DB 1 MSYQN -LVSEAGLITQKLLIHGDKEFLQHELKTIFARNWLEFLTHDSLIPAPGDVYTAAMG 58  
QY 61 IDEVIVSRONQSGIRAFINVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELOSVPFE 120  
DB 59 IDEVIVSRONQSGIRAFINVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELOSVPFE 118

QY 121 KDLYGESLNKKKCLGKEVARVESFHGFYTCGFDDEAPPLMDYLDGDAAMYLEPMPKHSGL 180  
DB 119 KELVGDATKKKCLGKEVARVESFHGFYTCGFDDEAPPLMDYLDGDAAMYLEPMPKHSGL 178

QY 181 ELVGPFGKVVYIKAMKKAPENFVGDYHVGTWTHASLSRSGESIFSSLAGNALPPEGAGL 240  
DB 179 ELVGPFGKVVYIKAMKKAPENFVGDYHVGTWTHASLSRSGESIFSSLAGNALPPEGAGL 238

QY 241 QMTSKYSGMGVLMDGYSVSHADLVPELMAFGAKQERLKEIGDVARIYRSHLNTCV 300  
DB 239 QMTSKYSGMGVLMDGYSVSHADLVPELMAFGAKQERLKEIGDVARIYRSHLNTCV 298

QY 301 FPNNSMLTCSGVFWVWNPIDANTTEVMTYATVEKMPEDLKRRLADSVQRTIGPAGFMS 360  
DB 299 FPNNSMLTCSGVFWVWNPIDANTTEVMTYATVEKMPEDLKRRLADSVQRTIGPAGFMS 358

QY 361 DDNDNMTASONGKRYOSRDSLSNLGFGEDVYGDVAVPGVKSALIGETSYRGFYRAY 420  
DB 359 DDNDNMTASONGKRYOSRDSLSNLGFGEDVYGDVAVPGVKSALIGETSYRGFYRAY 418

QY 421 QAHVSSNMWAEFEHASTWHTLTKTTDR 449  
DB 419 QAHVSSNMWAEFEHASTWHTLTKTTDR 447

## RESULT 6

T31134  
naphthalene dioxygenase (EC 1.14.12.-) large chain - Sphingomonas aromaticivorans pla  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31134  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom  
A:Reference number: Z20992  
A:Accession: T31134  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-459 <ROM>  
A:Cross-references: EMBL:AF079317; NID:93378261; PID:93378275; PIDN:AA003858.1  
C:Genetics:  
A:Gene: bphA1f  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:80,82,100,103/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 39.7%; Score 957; DB 2; Length 459;  
Best Local Similarity 43.9%; Pred. No. 1.6e-68;  
Matches 192; Conservative 71; Mismatches 162; Indels 12; Gaps 7;

QY 8 LVSESGLSQKHLIHGDEELFQHELKTIFARNWLEFLTHDSLIPAPGDVYTAAMGIDEVIVS 67  
DB 7 LVDNANASQSRVFWDDVQYQLELIFRSCWMLGDSLVPRKGDPIITTYMAEDRVILS 66

QY 68 RQNDGSRATLANVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELOSVPFEKDLGYES 127  
DB 67 RQPDGSRATLANVCRRHGRKTLVSEAGNAKGFVCSYHGWGSGNGELOSVPFEKDLGYES 126

QY 128 LNNKCLGKKEVARVESFHGFYTCGFDDEAPPLMDYLDGDAAMYLEPMPK -HSGGLTGVGP 186  
DB 127 LDKSKGLAPV -RVEYTKGIFGCHPEARSLEDYLDGDPWYLDITWDGDDGELIGLGP 185

QY 187 GKVVYIKAMKKAPENFVGDYHVGTWTHASLSRSGESIFSSLAGNALPPEGAGLQMTSK 245  
DB 187 GKVVYIKAMKKAPENFVGDYHVGTWTHASLSRSGESIFSSLAGNALPPEGAGLQMTSK 245

Db 186 LKSTLACNMKVPTEFNVGDYGVHGWTHAALOMIGELAGLSGNRADMPDGLIOETMR 245  
QY 246 YGSGMGVLMDSYGVH--SADLVPBELMAFGAKOERLNKEIGDVRARIYSHLCTEFPN 303  
Db 246 HGGFGGLIDNAATAIHKRQGYVLYLETRGRIKREKPE----RERLYGHMNTSLPN 301  
QY 304 NSMLTSGVGVKVPNPIDANTTEVWTVAIVEKDPEDLKRRLADSVQRTIGPFWESDDN 363  
Db 302 CSFLYGTNFKIMHPRPHIEVWTYMPVKNDTEKRSLQREAIRSPFGAGTLESDDG 361  
QY 364 DNNETASQNGKTKYQSRDLSNLGEGEDYVG--DAYPPGVGKSAIETSTRGFYAYQA 422  
Db 362 ENNSSATYNNNGGIITRKGRM--NSSMKDREGPHVPYGVIVGSFICETSYRGFYREWQE 419  
QY 423 HVSSSMAEFEHASTW 439  
Db 420 MLDAPDMAIRANDTW 436

RESULT 7  
T31256  
terminal oxygenase component large chain homoloy - Sphingomonas aromaticivorans plasmid  
C:Species: Sphingomonas aromaticivorans  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000  
C:Accession: T31256  
R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; G  
submitted to the EMBL Data Library, July 1998  
A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromati  
A:Reference number: Z20992  
A:Accession: T31256  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-450 <ROM>  
A:Cross-references: EMBL:AF079317; NID:93378261; PID:93378397; PIDN:AD03980.1  
C:Genetics:  
A:Genome: plasmid PNL1  
A:Note: bphA1a

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe  
C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein  
F:81,83,101,104/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 32.1%; Score 774; DB 2; Length 450;

Best local Similarity 36.1%; Pred. No. 6; 6e-54;

Matches 157; Conservative 88; Mismatches 184; Indels 6; Gaps 6;

QY 1 MNVNNKTIIVSEGLSOKHLHGDELEFQHEKTFARNWLFTHDSLIPAPGDVYTKMG 60  
Db 1 MNDSTADLVDSRTGROSRSIYASEDITROELERIFGCMLEFLVHTSOIPKGFDFRTFMG 60  
QY 61 IDEVIVSRONDGSIKAFILANVCRRHGTILVSEVGNANAFVCSYHGMGFGSNGELQSVPE 120  
Db 61 EDDVIVIRQKDSIKAFILNSCTHRGNRICRADGNARAFICNHYGMSFSPGALSGVPLE 120  
QY 121 KDLVGSLLNKKCLGLKEVAVESFHFYICGFDQEARPLMDYLGDAAWYLEP--MFKISGG 179  
Db 121 NEAYFELDRTKFGLIPVTKVAEYKGLVFGCWDANSPSLDYLGDAAEFPLDVLWDAMPGG 180  
QY 180 LELVGGPKVVIKANKKAPAEVNGDAYHVGWTHASSLSRGSEI--FESLA--GNALPREG 237  
Db 181 SALIGETQKAVLGTNNKLPVENVCSDGCTHIGMAHAGMAAAVQSDLTGLVSGNSGVLDG 240  
QY 238 AGLOMTSKYSGMGVLMDSYGVH--SADLVPBELMAFGAKOERLNKEIGDVRAR--IYRSHL 296  
Db 241 -GLSVAGMGHVALSLADGVSGYAFYDPDKPILEYLEANNGQVYIDRAGEVRGRQVMAQV 299  
QY 297 NCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDPEDLKRRLADSVQRTIGPAG 356  
Db 300 NITIFPNLLDLPGLNFRVYVHPKPGQIEQMTAMAMENDPEAVKAOILNEOCLTFGLG 359  
QY 357 FNESSDNDMMETASQNGKTKYQSRDLSNLGEGEDYVGAYPPGVGKSAIETSTRGF 416  
Db 360 LFDNDGDGDLTACTEQRGWRITQMDVYTNMALGRSGKREG--FPGDIAAGLVSEHNORYF 418

QY 417 YRAYOAHVSSSMNAE 431  
Db 419 YRRQEHMAETMAE 433

RESULT 8  
A65031  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escheri  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
C:Accession: A65031; S49292  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.: Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: A65031  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Molecule type: DNA  
A:Cross-references: GB:AE000340; GB:U00096; NID:91788883; PIDN:AC75591.1; PID:917888  
A:Experimental source: strain K-12, substrain MG1655  
R:Turin, E.; Gasser, F.; Biville, F.  
Submitted to the EMBL Data Library, September 1994  
A:Description: Cloning and sequencing of an E. coli gene homologous to dioxygenase of  
A:Reference number: S49292  
A:Accession: S49292  
A:Molecule type: DNA  
A:Residues: 1-119, 'A', 21-383, 'ATAPATANCWKW', 397-398, 451, 'R', 453, 'SAATATLALITTSFQKPL  
A:Cross-references: EMBL:Z37966; NID:9350595; PIDN:CA86018.1; PID:9350596  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2  
C:Keywords: 2Fe-2S; metalloprotein; oxidoreductase; Rieske iron-sulfur protein  
F:85,87,105,108/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 30.6%; Score 737; DB 2; Length 453;

Best local Similarity 36.4%; Pred. No. 6e-51;

Matches 165; Conservative 64; Mismatches 174; Indels 50; Gaps 8;

QY 20 IHGDELEFQHEKTFARNWLFTHDSLIPAPGDVYTKMGIDEVIVSRONGSIRAFIN 79  
Db 24 IYDPDILYOLEKERTIFRCWLFANESQIPKPDFFNTYGEDAVVYVVKQKDSIRAFIN 83  
QY 80 VCRHGGKTLVSEVGNANAFVCSYHGMGFGSNGELQSVPEKDLVGEISNKKCLGLKEVA 139  
Db 84 QCRHRAKRVSYADCGNTRAFETCPYHGWSYGINGELIDVLEPRAYPQGLCKSHWGLNEVP 143  
QY 140 RVESFHFYICGFDQEARPLMDYLGDAAWYLEP--MFKISGGLELVCPGKVIKANKKAP 198  
Db 144 CVESYKLLIFGNMDSAPGLRDYLDIAWYLDMLDRREGTEIVGVQKWIINCMKPEP 203  
QY 199 AENFVGAYHVGWTHASSL-----RSGSEIFSSLAGNALP----- 234  
Db 204 AEOFASQYIHALFSHASAVOYLAKKDDGSKRLD-----GQTRAPVWETAKDALQFG 256  
QY 235 --PEGAGLOMTSKYSGMGVLMDSYGVH--SADLVPBELMAFGAKOERLNKEIGDVRARIY 292  
Db 257 QDGGSGSEFFTEKPKDANVWV--DGAVSYRYETVAE-----AQGRLEGVARAL 303  
QY 293 RSHLINCTVFPNNSMLTCSGVFKVWNPIDANTTEVWTVAIVEKDPEDLKRRLADSVQRTI 352  
Db 304 AGHANN--IFPILLSLANTATLRVWHPRGPQVYVMAFCITDKAASDEYKAAEFNSATRAF 361  
QY 353 GPAGFESDDNDMMETASQNGKTKYQSRDLSNLGEGEDYVDAYPPGVGKSAIETSTRGF 412  
Db 362 GPAGFELDQDSENNCTQIKLKGRHANSKLCLEMGLGCKRRDDIGIT--NYIFSETA 420  
QY 413 YRGFYRAYOAHVSSSMNAEFEHASTWHTELTK 445  
Db 421 ARGMYQWADLLSSESWQEVLDKTAAYVQCEYMK 453

RESULT 9

D91054  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 31-Dec-2001  
C:Accession: D91054  
R:Hayashi, T., Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasasavar, N.; Yasunaga, T.; Kunara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic islands  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: D91054  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-453 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BAR36827.1; PID:G13362875; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain R1MD 0509952  
A:Genetics:  
A:Gene: ECs3404  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe-2S] ferredoxin  
:Keywords: oxidoreductase

Query Match	30.6%;	Score 737;	DB 2;	Length 453;
Best Local Similarity	36.4%;	Pred. No. 6e-51;		
Matches 165;	Conservative 64;	Mismatches 174;	Indels 50;	Gaps 8;

```

QY 20 IHGDEEFOHBLKTIFFRNMWLEFTHDSLIPAGQVYPAKMGIDEIVSRONDGSIIRAFIN 79
Db 24 IYTPDIQOLEBETIFGCWMLFLAHESQIPKRGDFNNYMGEDAVVYVRQKDSIKAFIN 83
QY 80 VCRHGRKTVLSVEANGAKGEVCSYHGMWFGSGNGELQSVPEKDLYGESLKKCLGLEVA 139
Db 84 QCRRRAMVSYADCGNIRAFPCPYHWSYGINGELIDPLEPRAYPOGLCKSHMGLINEVP 143
QY 140 RVESFHGTYICCEQOEARPLMDYLGDAWYLEPME-KHSGLELYVPRGKYIANKAP 198
Db 144 CVESIKGLIFGNWPTSAAGLRDYLGDIAWYLDGMLDREGGTETIVGYQKWVINCNMKEP 203
QY 199 AENEVDGAYHYHGWTHASL-----RSGESIFSSLAGNAALP----- 234
Db 204 AEQFASPDQYHALFHASAVQYLAKDQSGDKRLD-----GQFARVWETAKDALQFG 256
QY 235 --PBGAGLOMTSKTSGSGMGVLMGDSGVYHSADLYPELMAGFGAKOERLKEIGVRIY 292
Db 257 QDGHGSGFFFEKEDANWV--DGAVSYYRETYAE-----AEQRLGEVRLRL 303
QY 293 RSHNCGYFEPNNSMLTCSGVGEKVMNPIDANTTEWTYAIYKEMPEDKRLRLASVQRTI 352
Db 304 AGHNN--IFPLSLMNGTATLRVHMPRGPDQVEYWAFCITDKAASDEKKAFAESATPAF 361
QY 353 GPAGFWESEDNDNNETASQNGKQYQSRSDLSNLGEGEDVYGAUVPYGVYKSAIGETS 412
Db 362 GPAGFLEQDDSENNCEIQIKLLKGHRARNSKICLEMGLGQEKRRDQDIPGL-NTYFSETA 420
QY 413 YRGFYRAYQAHVSSNNAEFPHASSTWHTELTK 445
Db 421 ARGMYORWADLLSESWQEVLDKTTAAVQOEYVK 453

```

RESULT 10  
G85898  
biphenyl dioxygenase (EC 1.14.-.-) terminal oxygenase component large chain - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Dec-2001  
C:Accession: G85898  
R:Pena, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamoustis, K.; Apodaca  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G85898  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-453 <STO>

A:Cross-references: GB:AE005174; NID:912516944; PIDN:AAG57651.1; GSPDB:GN00145; UMG:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: hcaal  
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [4]  
C:Keywords: oxidoreductase

Query Match	30.68;	Score 737;	DB 2;	Length 453;
Best Local Similarity	36.48;	Pred. No. 6e-51;		
Matches 165;	Conservative 64;	Mismatches 174;	Indels 50;	Gaps 8;

```

0Y 20 IHGDELEFOHEKTEFTFANFWLFLTHDSLIPAPGDVVTAKGIDEIVYSRONDGSTRAFLN 79
Db 24 IYTPDDIYQLELERTFGCWLFLAHESQIPRPGDFPNITMGEDAVVYVRKODGSTKAELN 83
0Y 80 VCRHRKTLVSVEAGNAKGVCOSYHGWGFGSNGELOSVPEKDLVGESLNKKCLGLEVA 139
Db 84 QCRHRAMRVSYADCGCNTBATFCPTYHGMYSYGINGELIDVPLPRAYPOGLCKSHMGLNEVP 143
0Y 140 RVESPHGFTYCFOEDAPRLMDYLGDAWYLEBPMF-KHSGGLEIVPGPKVIVANKMAP 198
Db 144 CVESKTKGLFGWMDTSAGLDYLDYLDIAMDYLDGMIDRREGSTELYGVQKWNINCMKFP 203
0Y 199 AENFVGDAYVHYHGWTHASSL-----RSGSEIFSSLAGNAALP----- 234
Db 204 AEOFASDQYHALFHSASAVQYLAKDDGSDRKLTD-----GQARVWMTAKDALQFG 256
0Y 235 --PBAGACLOMTSKSGSGKGYLMDYISGVHSAADYVPELMAFGAKOERLNKETIDVRAIY 292
Db 257 QDGHOSGFFFEKPRDANVW--DGAVSSYRETYAE-----AEORIGEVRAURL 303
0Y 293 RSHLNCYTFEPNNSMITCSGVFKVWMPIDANTTEVYTAIVEKMPDEDLKRLRADSVRTI 352
Db 304 AGHNH--IFPLLSWLNGLATLRTVWHPRGPDQVEYWAFCITLTKASDEYKAAPENSATAPF 361
0Y 353 GPAGFWESDDDNNMETASQNGKKYOSRDSLLSNLGFCEGDYVGAUVYPGVYKSAIGETS 412
Db 362 GPAGLEODDESENMCIOIKLLKGRHARNSKLCLEWGLGOEKRDRDGPICIT--NYIFSETA 420
0Y 413 YRGFYRAVOAHVSSSNMAFEHASTWHTLTK 445
Db 421 ARGMTYORMADLLSSESQEWELDKTAYAOOEYWK 453

```

RESULT 11

T31258

aromatic oxygenase large chain - Sphingomonas aromaticivorans plasmid pNL1

C:Species: Sphingomonas aromaticivorans

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 19-May-2000

C:Accession: T31258

R:Römling, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W. submitted to the EMBL Data Library, July 1998

A:Description: Complete sequence of a 164 kb catabolic plasmid from Sphingomonas arom

A:Reference number: Z20992

A:Accession: T31258

A:Status: preliminary;

A:Molecule type: DNA

A:Residues: 1-455 <ROM>

A:Cross-references: EMBL:AF079317; NID:g3378261; PTD:g3378399; PIDN:APD03982.1

C:Genetics:

A:Genome: plasmid pNL1

A:Note: bphA1b

C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [

C:Keywords: 2Fe-2S; metalloprotein; Rieske iron-sulfur protein

F:88,90,108,111/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 30.1%; Score 725.5; DB 2; Length 455;

Best Local Similarity 36.1%; Pred. No. 5e-50;

Matches 153; Conservative % 83; Mismatches 169; Indels 19; Gaps 8;

0y 2 NNNKKILVSEGSOKHLIHGDEELFQHELKTTRARNMLFLTHSLSLPAPDYYTAKMGI 61

db 11 DYSRVMDKEDWLDNR--IFSDADYIEELRIFARSWLFAVHESQIPSSGDFLTTHNGE 68

```
OY 62 DEVIYSRONDGSIKRAFLNVRHGRKTLVSVEAGNAKGFVCSYHGMGFGSNGELQSVPEK 121
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 69 DAVIVANQPGSIRVYMLNSCPHRGNKVCYFADAGNTRRVCYCHOMAFPTADLGMHEY 128
OY 122 DLYGESLNKCKLGEKAVARVESFHGTYGCFDQAPPLMDYLGAAMYLEPMF-KHSGCL 180
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 129 CYDAGDIDFKNHGLKNVAKVGNVKGVLFAFENSDAPSLAELGDFWRLDMLILNEEGE 188
OY 181 ELVGPPEKVVYIKANWKPAENFVGDAYHVGTTHASLSRSGSIFSSLAGNALPPEGGL 240
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 189 EFTGGCTKSYISANWKGVENFIDGDAYHAGWTHDSGTR-----SMNNGCPFPF-----I 237
OY 241 QMTSKYSGM-GVLM-D-GYSVHSADLV--PELMAFGAKOERLNKEIGDVRARIYRSHL 296
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 238 DMONSXYHASYNGHGMERGTEGVDLFLGRPKVMDYVNIKIRPKMAERLGEKRSKIFGSA 297
OY 297 NCIVFPNKSMLTCSGVKVNPNIDANTTEWVTYAIVEKMPEDLKRRLADSVQRTIGPAG 356
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 298 SASIFPVSFLPGISTFRQWPKGPMQELKTWIVNKNMDDIKEEVTKGVMQTFGPGG 357
OY 357 FWSDDMDNNETASQNGKTKYQSRDLSNLGFGEDYVGDVYVGVYKSAIGENYRGF 416
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 358 TFMDDGEMNENCTTVNRGVYTRHERLHYRCIGRIQIEHDTL-PGIYIRGOYDANORGF 416
OY 417 YRAY 420
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 417 YQRM 420

RESULT 12
JN0812
benzene 1,2-dioxygenase (EC 1.14.12.3) terminal oxygenase component large chain - Pseudoc
C:Species: Pseudomonas putida
C>Date: 19-May-1994 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: JN0812
R:Itan, H.M.; Tang, H.Y.; Joannou, C.L.; Abdel-Wahab, N.H.; Mason, J.R.
Gene 130, 33-39, 1993
A:Title: The Pseudomonas putida ML2 plasmid-encoded genes for benzene dioxygenase are un
A:Reference number: JN0810; MUID:9345820; PMID:8344526
A:Accession: JN0812
A:Molecule type: DNA
A:Residues: 1-450 <ZVL>
A:Cross-references: GB:L04642; GB:L04643; NID:96552505; PID:AAA17758.1; PID:9309855
A:Experimental source: strain ML2
C:Comment: This enzyme is involved in catalyzing the oxidation of benzene to cis-1,2-dih
C:Genetics: The reduced enzyme catalyzes the oxidation of benzene to cis-benzene dihydrod
A:Gene: dedc1
A:Genome: plasmid
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske [2Fe
C:Keywords: 2Fe-2S; aromatic hydrocarbon catabolism; metalloprotein; oxidoreductase; Rie
F:86-134/Domain: Rieske [2Fe-2S] homology <RSK>
F:96,98,116,119/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status pre

Query Match 27.1%; Score 652.5; DB 1; Length 450;
Best Local Similarity 34.6%; Pred. No. 3,4e-44;
Matches 148; Conservative 69; Mismatches 174; Indels 37; Gaps 11;

OY 20 IHGDEELFOHELKTIFARNMLFLTHDSLIPAPGQYTAKKGIDIEVIYSRONDSIRAFIN 79
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 35 IYDEDLYOLELERRVFAKRWMLLGHETHIRKPGDYFTTYGDEPVVYVVRKODASIAVFLN 94
OY 80 VCRHGRKTLVSVEAGNAKGFVCSYHGMGFGSNGELQSVPEKDLYGESLNKCKLGE-- 137
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 95 QCHRHGRICRADAGNAKATTCYHGWAYDTAGNLINVPYEAESFA-----CLDKKEMS 148
OY 138 --VARVESFHGTYGCFDQAPPLMDYLGAAMYLEPMF-KHSGCLVGVPPGKVYIKAN 194
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 149 PLKARVETVYGLIFANMDENAVLDLDTYLGEAKFYMDHMDRTEAGTEVTEIQLKWIIPCN 208
OY 195 WKPAENFVGDAYHVGMT-HASSLRSGESIFSSLAGNALPPEGAGLOMTSKYSGMGVL 253
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```

```
Db 209 WKFAEOPCSDMHNAGTTAHLGIIAGLPBDELADLA--P-----KFGKOYRAS 257
OY 254 WDGY-SGVHSHAD-----LVPELMAF--GAKOERLNKEIGDVR--ARIYRSHLCTVFE 301
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 258 WGHGSGFYIDDPMLMLAMGPKYTSTLTEGPAKAEKAEKLGISTEKGTKMLEHM--TYF 315
OY 302 PNNMSLTCSGVFVWNPIDANTTEWVTYAIVEKMPEDLKRRLADSVQRTIGPAGWESD 361
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 316 PTCSEFLPQVNTIRTMHPRGRNEVEVMAFTVVDADAPDIDKEEPRRQTLRTFSAGVFEED 375
OY 362 DNDMNETASONGKTKYQSRDLSNLGFGEDYVGDVYVGVYKSAIGENYRGFRAAO 421
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 376 DGMWVEIQLTHLRGHKARSRPFNAMSGQTVNDPIYGRISNNVYSEEAARGLYAHWL 435
OY 422 AHVSSSNW 429
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 436 KMTSPDM 443

RESULT 13
A36516
toluene dioxygenase (EC 1.14.12.11) terminal oxygenase component large chain - Pseudo
C:Species: Pseudomonas putida
C>Date: 15-Feb-1991 #sequence_revision 06-Jan-1995 #text_change 18-Jun-1999
C:Accession: A36516
R:Zylstra, G.J.; Gibson, D.T.
J. Biol. Chem. 264, 14940-14946, 1989
A:Title: Toluene degradation by Pseudomonas putida F1. Nucleotide sequence of the tod
A:Reference number: A36516; MUID:89359301; PMID:2670929
A:Accession: A36516
A:Molecule type: DNA
A:Residues: 1-450 <ZVL>
A:Cross-references: GB:J04996; NID:g151600; PIDN:AAA26005.1; PID:g151601
A:Experimental source: strain F1
C:Genetics:
A:Gene: todC1
C:Superfamily: toluene dioxygenase terminal oxygenase component large chain; Rieske I
C:Keywords: 2Fe-2S; aromatic hydrocarbon catabolism; metalloprotein; oxidoreductase;
F:86-134/Domain: Rieske [2Fe-2S] homology <RSK>
F:96,98,116,119/Binding site: 2Fe-2S cluster (Cys, His, Cys, His) (covalent) #status

Query Match 27.0%; Score 650.5; DB 1; Length 450;
Best Local Similarity 34.4%; Pred. No. 4,9e-44;
Matches 145; Conservative 71; Mismatches 181; Indels 25; Gaps 9;

OY 20 IHGDEELFOHELKTIFARNMLFLTHDSLIPAPGQYTAKKGIDIEVIYSRONDSIRAFIN 79
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 35 IYDEDLYOLELERRVFAKRWMLLGHETHIRKPGDYFTTYGDEPVVYVVRKODASIAVFLN 94
OY 80 VCRHGRKTLVSVEAGNAKGFVCSYHGMGFGSNGELQSVPEKDLYGESLNKCKLGE-- 137
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 95 QCHRHGRICRADAGNAKATTCYHGWAYDTAGNLINVPYEAESFA-----CLNKEMS 148
OY 138 --VARVESFHGTYGCFDQAPPLMDYLGAAMYLEPMF-KHSGCLVGVPPGKVYIKAN 194
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 149 PLKARVETVYGLIFANMDENAVLDLDTYLGEAKFYMDHMDRTEAGTEVTEIQLKWIIPCN 208
OY 195 WKPAENFVGDAYHVGMT-HASSLRSG--ESIFSSLAGNALPPEGAGLOMTSKYSGMG 251
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 209 WKFAEOPCSDMHNAGTTSHLGLAGLPEDL-----EMADLPPIYVGVKQYRASMGHGS 263
OY 252 VLMDYSGVSHADLVPELMAF--GAKOERLNKEIGDVR--ARIYRSHLCTVFPNNSM 307
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 264 GFVYGDPLMLALMIAIGPVTYVTEGPAKAEKAEKLGISTEKGTKMLEHM--TYFPCSTL 321
OY 308 TCSGVFVWNPIDANTTEWVTYAIVEKMPEDLKRRLADSVQRTIGPAGWESDNDNNE 367
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 322 PGITVTRTWHPRGRNEVEVMAFTVVDADAPDIDKEEPRRQTLRTFSAGVFEEDDENNV 381
OY 368 TASONGKTKYQSRDLSNLGFGEDYVGDVYVGVYKSAIGENYRGFRAAYQAHVSS 427
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 382 EIOHILRGHARSFPFNAMSGQTVNDPIYGRISNNVYSEEAARGLYAHHLRMWISP 441
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
```







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 08:46:47 : Search time 29.1667 Seconds  
(without alignments)  
2051.298 Million cell updates/sec

Title: US-09-843-250-36  
Perfect score: 2408  
Sequence: 1 MNVNNKILVSESGLSQKHLI.....AEFEHASTWHTLTKTTDR 449

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2408	100.0	449	21	AA12582
2	2407	100.0	449	21	AA12585
3	2407	100.0	449	21	AA12586
4	2407	100.0	449	21	AA12587
5	2406	99.9	449	21	AA12581
6	2404	99.8	449	21	AA12577
7	2403	99.8	449	21	AA12579
8	2403	99.8	449	21	AA12580
9	2402	99.8	449	21	AA12586
10	2401	99.7	449	21	AA12583

11	2400	99.7	449	21	AA12578
12	2398	99.6	448	21	AA12584
13	2397	99.5	449	21	AA12569
14	2349	97.5	449	21	AA12570
15	2321	96.4	449	21	AA12571
16	2295	95.3	449	21	AA12572
17	2222	92.3	449	21	AA12573
18	2192	91.0	447	21	AA12574
19	2055	85.3	447	21	AA12575
20	1971.5	81.9	451	21	AA12576
21	1095	45.5	452	21	AA12577
22	737	30.6	453	22	AA12578
23	661.5	27.5	458	21	AA12579
24	657.5	27.3	443	23	AA12580
25	623.5	25.9	459	16	AA12581
26	618	25.7	458	21	AA12582
27	615	25.5	458	21	AA12583
28	554.5	23.0	427	14	AA12584
29	377	15.7	490	22	AA12585
30	376.5	15.6	385	22	AA12586
31	358	14.9	497	22	AA12587
32	341.5	14.2	424	22	AA12588
33	239	9.9	435	22	AA12589
34	235	9.8	433	22	AA12590
35	235	9.8	433	22	AA12591
36	233	9.7	438	23	AA12592
37	224	9.3	446	19	AA12593
38	223	9.3	439	19	AA12594
39	196	8.1	99	23	AA12595
40	186	7.7	426	23	AA12596
41	184	7.6	1193	22	AA12597
42	181	7.5	35	19	AA12598
43	173	7.2	35	19	AA12599
44	164	6.8	35	19	AA12600
45	126	5.2	405	19	AA12601

#### ALIGNMENTS

RESULT 1	
AA12582	
ID	AA12582 standard; Protein; 449 AA.
AC	AA12582;
DT	09-NOV-2000 (first entry)
DE	Naphthalene dioxygenase mutant F352I protein sequence SEQ ID NO:36.
OS	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO.
OS	Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KW	polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX	
XX	Pseudomonas sp.
OS	Synthetic.
PN	WO200037480-A1.
PD	29-JUN-2000.
PF	26-OCT-1999; 99WO-US25079.
PR	26-OCT-1998; 98US-0105575.
PA	(IOWA) UNIV IOWA RES FOUND.
PI	Parales R, Gibson D, Resnick S, Lee K;
DR	WPI: 2000-452174/39.
DR	N-PSDB: AAA65356.
PT	Novel naphthalene dioxygenase mutant having a specific amino acid

PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PS  
XX  
PS Claim 9; Page 133-134; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA;

Query Match 100.0%; Score 2408; DB 21; Length 449;

Best Local Similarity 100.0%; Pred. No. 1.2e-224; Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYNNKIIVSSEGLSQKHLHGDELFQHEKLTIFARWMLFTLHDSLIPAGDYVTAKMG 60  
DB 1 MNYNNKIIVSSEGLSQKHLHGDELFQHEKLTIFARWMLFTLHDSLIPAGDYVTAKMG 60  
QY 61 IDEVIYSRONGSTRAFNLNVCRRHKGKTLVSEAGNAGFCVCSHGMGFSGNGLQSVPE 120  
DB 61 IDEVIYSRONGSTRAFNLNVCRRHKGKTLVSEAGNAGFCVCSHGMGFSGNGLQSVPE 120  
QY 121 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDAPPLMDYLGDAWYLEPMFKHSGGL 180  
DB 121 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDAPPLMDYLGDAWYLEPMFKHSGGL 180  
QY 181 ELVGPPEGKVVYKAMWKAPAEENFVGDAYHVGWTHASSLSRSGESIFSSLAGNALPREGAGL 240  
DB 181 ELVGPPEGKVVYKAMWKAPAEENFVGDAYHVGWTHASSLSRSGESIFSSLAGNALPREGAGL 240  
QY 241 QMTSKYSSGMVLMDGYSGVSHADLVPFLMAFGAKOERLNKEIGDVARIRSLNLCV 300  
DB 241 QMTSKYSSGMVLMDGYSGVSHADLVPFLMAFGAKOERLNKEIGDVARIRSLNLCV 300  
QY 301 FPNNSMILTCSCGVFKVWNPIDANTTEVMYTAIVEXKMPEDLKRRLADSVORTIGPAGFMS 360  
DB 301 FPNNSMILTCSCGVFKVWNPIDANTTEVMYTAIVEXKMPEDLKRRLADSVORTIGPAGFMS 360  
QY 361 DDNDNMETASONGKRRYOSRSDLLSNLGFGEDEVYGDVAYPGVVGSAIGETSYRGFTYRAY 420  
DB 361 DDNDNMETASONGKRRYOSRSDLLSNLGFGEDEVYGDVAYPGVVGSAIGETSYRGFTYRAY 420  
QY 421 QAHVSSSNMAEFEHASSTWHTELTKTTDR 449  
DB 421 QAHVSSSNMAEFEHASSTWHTELTKTTDR 449

RESULT 2  
AAB12565  
ID AAB12565 standard; Protein: 449 AA.  
XX  
AC AAB12565;

XX 09-NOV-2000 (first entry)

DE Naphthalene dioxygenase mutant F352V protein sequence SEQ ID NO:2.

KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.

OS Pseudomonas sp.  
OS Synthetic.

PN WO200037480-A1.

XX 29-JUN-2000.

PF 26-OCT-1999; 99WO-US25079.

XX 26-OCT-1998; 98US-0105575.

PA (IOWA ) UNIV IOWA RES FOUND.

PI Parales R, Gibson D, Resnick S, Lee K;

DR WPI: 2000-452174/39.

DR N-PSDB; AAA65339.

PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT

PS Claim 9; Page 58-60; 151pp; English.

CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA;

Query Match 100.0%; Score 2407; DB 21; Length 449;

Best Local Similarity 99.8%; Pred. No. 1.5e-224; Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNYNNKIIVSSEGLSQKHLHGDELFQHEKLTIFARWMLFTLHDSLIPAGDYVTAKMG 60  
DB 1 MNYNNKIIVSSEGLSQKHLHGDELFQHEKLTIFARWMLFTLHDSLIPAGDYVTAKMG 60  
QY 61 IDEVIYSRONGSTRAFNLNVCRRHKGKTLVSEAGNAGFCVCSHGMGFSGNGLQSVPE 120  
DB 61 IDEVIYSRONGSTRAFNLNVCRRHKGKTLVSEAGNAGFCVCSHGMGFSGNGLQSVPE 120  
QY 121 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDAPPLMDYLGDAWYLEPMFKHSGGL 180  
DB 121 KDLYGESLNKKCLGKEVARVESFHGFTYGCDFQDAPPLMDYLGDAWYLEPMFKHSGGL 180

Db 121 KDLGESLNKKCLGLKEVARVESFHGFIYGCFDQEAAPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPFGKVIKANKKAPAEFNVGDAYHYGWTTHASSLRGSEISFSSLAGNALPPEGAGL 240  
 Db 181 ELVGPFGKVIKANKKAPAEFNVGDAYHYGWTTHASSLRGSEISFSSLAGNALPPEGAGL 240  
 QY 241 QMTSKYXGSGMGLMDGYSGVSHADLVPPELMARFGAKQERLNKEIGDVRARIYRSHLNCYV 300  
 Db 241 QMTSKYXGSGMGLMDGYSGVSHADLVPPELMARFGAKQERLNKEIGDVRARIYRSHLNCYV 300  
 QY 301 FPNNSMLTCSGVFKYWNPIPDANTTEWMTAIVKMDPEDLKRRLLDSVORTGPGAFWES 360  
 Db 301 FPNNSMLTCSGVFKYWNPIPDANTTEWMTAIVKMDPEDLKRRLLDSVORTGPGAFWES 360  
 QY 361 DDNDMMETASQNGKKYQSRDLSNLGFGEDVYGDAPVGVGSAIGETSYRGFYRAY 420  
 Db 361 DDNDMMETASQNGKKYQSRDLSNLGFGEDVYGDAPVGVGSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449  
 Db 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449  
 RESULT 3  
 AAB12566  
 ID AAB12566 standard; Protein: 449 AA.  
 AC AAB12566;  
 DT 09-NOV-2000 (first entry)  
 DE NDO related complex alpha subunit protein sequence SEQ ID NO:14.  
 XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 OS Pseudomonas sp.  
 OS Synthetic.  
 PN WO200037480-A1.  
 XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 PE 26-OCT-1998; 98US-0105575.  
 PR 26-OCT-1998; 98US-0105575.  
 PA (IOWA ) UNIV IOWA RES FOUND.  
 PI Parales R, Gibson D, Resnick S, Lee K;  
 DR WPI: 2000-452174/39.  
 DR N-PSDB: AAA65340.  
 XX Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT -  
 XX Claim 13; Page 99-100; 151pp; English.  
 PS The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (I) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which

CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothioephene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.  
 SQ Sequence 449 AA:  
 Query Match 100.0%; Score 2407; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred No. 1,5e-224;  
 Matches 448; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNYNKKILVSESGLSQKHLIHGDEELFQHELKTI FARNNLFLTHDSLIPAPGDYVYAKMG 60  
 Db 1 MNYNKKILVSESGLSQKHLIHGDEELFQHELKTI FARNNLFLTHDSLIPAPGDYVYAKMG 60  
 QY 61 IDEVIVSRONDGSI RFLNVCNHRGKTLVSVAGNAKGFVCSYHGWGFGSNGELOSVPPE 120  
 Db 61 IDEVIVSRONDGSI RFLNVCNHRGKTLVSVAGNAKGFVCSYHGWGFGSNGELOSVPPE 120  
 QY 121 KDLGESLNKKCLGLKEVARVESFHGFIYGCFDQEAAPLMDYLGDAAWYLEPMFKHSGGL 180  
 Db 121 KDLGESLNKKCLGLKEVARVESFHGFIYGCFDQEAAPLMDYLGDAAWYLEPMFKHSGGL 180  
 QY 181 ELVGPFGKVIKANKKAPAEFNVGDAYHYGWTTHASSLRGSEISFSSLAGNALPPEGAGL 240  
 Db 181 ELVGPFGKVIKANKKAPAEFNVGDAYHYGWTTHASSLRGSEISFSSLAGNALPPEGAGL 240  
 QY 241 QMTSKYXGSGMGLMDGYSGVSHADLVPPELMARFGAKQERLNKEIGDVRARIYRSHLNCYV 300  
 Db 241 QMTSKYXGSGMGLMDGYSGVSHADLVPPELMARFGAKQERLNKEIGDVRARIYRSHLNCYV 300  
 QY 301 FPNNSMLTCSGVFKYWNPIPDANTTEWMTAIVKMDPEDLKRRLLDSVORTGPGAFWES 360  
 Db 301 FPNNSMLTCSGVFKYWNPIPDANTTEWMTAIVKMDPEDLKRRLLDSVORTGPGAFWES 360  
 QY 361 DDNDMMETASQNGKKYQSRDLSNLGFGEDVYGDAPVGVGSAIGETSYRGFYRAY 420  
 Db 361 DDNDMMETASQNGKKYQSRDLSNLGFGEDVYGDAPVGVGSAIGETSYRGFYRAY 420  
 QY 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449  
 Db 421 QAHVSSSNMAEFERHASTWHTELTKTTDR 449  
 RESULT 4  
 AAB12567  
 ID AAB12567 standard; Protein: 449 AA.  
 AC AAB12567;  
 DT 09-NOV-2000 (first entry)  
 DE NDO related complex alpha subunit protein sequence SEQ ID NO:15.  
 XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
 KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
 KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
 OS Pseudomonas sp.  
 OS Synthetic.  
 PN WO200037480-A1.  
 XX 29-JUN-2000.  
 PD 26-OCT-1999; 99WO-US25079.  
 PR 26-OCT-1999; 99WO-US25079.



```

QY 1 MNYNNKILVSESGLSQKHILHGEDELFOHELKTIFFARMNMLFTHDSLIPAPGVYAKMG 60
DB 1 MNYNNKILVSESGLSQKHILHGEDELFOHELKTIFFARMNMLFTHDSLIPAPGVYAKMG 60
QY 61 IDEVIYSRONDGSIRAFPLANCRRKGTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120
DB 61 IDEVIYSRONDGSIRAFPLANCRRKGTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120
QY 121 KDLVGSLLNKKCLGLKEVAVRSEFHGFIYGCDFQDEAPPLMDYLGDAWYLEPFRKHSGL 180
DB 121 KDLVGSLLNKKCLGLKEVAVRSEFHGFIYGCDFQDEAPPLMDYLGDAWYLEPFRKHSGL 180
QY 181 ELVGPFGKYVIRKANMKAPAEVNFYGDAYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240
DB 181 ELVGPFGKYVIRKANMKAPAEVNFYGDAYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240
QY 241 QMTSKYSGMGVLMDSYGVSHADLVPBELMAFGAOKERLNKEIGVRRARIYSHLNCYV 300
DB 241 QMTSKYSGMGVLMDSYGVSHADLVPBELMAFGAOKERLNKEIGVRRARIYSHLNCYV 300
QY 301 FPNNMILTCSGVFKVWNPIDANTTEVMTYAIVEKMDPEDLKRRLADSVORTGPAGFWES 360
DB 301 FPNNMILTCSGVFKVWNPIDANTTEVMTYAIVEKMDPEDLKRRLADSVORTGPAGFWES 360
QY 361 DDNDNMETASONGKYYQSRDSDLNMLGFGEDYVGDVYPGVVGKSAIGETSYRGFYRAY 420
DB 361 DDNDNMETASONGKYYQSRDSDLNMLGFGEDYVGDVYPGVVGKSAIGETSYRGFYRAY 420
QY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449
DB 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

RESULT 6
AAB12577
ID AAB12577 standard. Protein: 449 AA.
XX
AC AAB12577;
XX
DT 09-NOV-2000 (first entry)
XX
DE Naphthalene dioxygenase alpha subunit protein sequence SEQ ID NO:26.
XX
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
XX polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX
OS Pseudomonas sp.
XX
PN WO200037480-A1.
XX
PD 29-JUN-2000.
XX
PE 26-OCT-1999; 99WO-US25079.
XX
PR 26-OCT-1998; 98US-0105575.
XX
PA (IOWA ) UNIV IOWA RES POUND.
XX
PI Parales R, Gibson D, Resnick S, Lee K;
XX
DR WPI, 2000-452174/39.
XX
DR N-PSDB; AAA65351.
XX
PT Novel naphthalene dioxygenase mutant having a specific amino acid
PT substitution for preparing chiral diols for use in the polymer, resin,
PT pharmaceutical or rubber industry and for carrying out bioremediation
XX
XX
XX Disclosure: Page 118-120; 151pp; English.
XX
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO

```

```

CC related complex (I) comprising several polypeptides which contain an
CC alpha subunit that contains substituted amino acids at specific
CC positions. The polypeptides and host cells are useful for preparing
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol (- or +)-cis biphenyl-3,4-
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
CC involves contacting them with naphthalene, biphenyl, phenanthrene,
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
CC involves contacting them with phenanthrene. The polypeptides and the
CC host cells are also used in bioremediation in which they oxidise an
CC aromatic compound such as indene, 1,2-dihydronaphthalene,
CC benzocyclopent-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
CC dibenzochiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a
CC corresponding dihydro dihydroxy compound. The polypeptides and the host
CC cells are useful for preparing chiral diols for use in the polymer,
CC resin, pharmaceutical or rubber industry. The present sequence represents
CC naphthalene dioxygenase (NDO) from the Pseudomonas sp. strain NCIB
CC 9816-4, which is used in the exemplification of the present invention.
XX
SQ Sequence 449 AA:
XX
Query Match 99.8%; Score 2404; DB 21; Length 449;
Best Local Similarity 99.8%; Pred. No. 2.9e-224;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MNYNNKILVSESGLSQKHILHGEDELFOHELKTIFFARMNMLFTHDSLIPAPGVYAKMG 60
DB 1 MNYNNKILVSESGLSQKHILHGEDELFOHELKTIFFARMNMLFTHDSLIPAPGVYAKMG 60
QY 61 IDEVIYSRONDGSIRAFPLANCRRKGTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120
DB 61 IDEVIYSRONDGSIRAFPLANCRRKGTILVSEAGNAKGFVCSYHGMFGSNGELQSVPE 120
QY 121 KDLVGSLLNKKCLGLKEVAVRSEFHGFIYGCDFQDEAPPLMDYLGDAWYLEPFRKHSGL 180
DB 121 KDLVGSLLNKKCLGLKEVAVRSEFHGFIYGCDFQDEAPPLMDYLGDAWYLEPFRKHSGL 180
QY 181 ELVGPFGKYVIRKANMKAPAEVNFYGDAYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240
DB 181 ELVGPFGKYVIRKANMKAPAEVNFYGDAYHVGWTHASSLRSGESIFSSLAGNAALPPEGAGL 240
QY 241 QMTSKYSGMGVLMDSYGVSHADLVPBELMAFGAOKERLNKEIGVRRARIYSHLNCYV 300
DB 241 QMTSKYSGMGVLMDSYGVSHADLVPBELMAFGAOKERLNKEIGVRRARIYSHLNCYV 300
QY 301 FPNNMILTCSGVFKVWNPIDANTTEVMTYAIVEKMDPEDLKRRLADSVORTGPAGFWES 360
DB 301 FPNNMILTCSGVFKVWNPIDANTTEVMTYAIVEKMDPEDLKRRLADSVORTGPAGFWES 360
QY 361 DDNDNMETASONGKYYQSRDSDLNMLGFGEDYVGDVYPGVVGKSAIGETSYRGFYRAY 420
DB 361 DDNDNMETASONGKYYQSRDSDLNMLGFGEDYVGDVYPGVVGKSAIGETSYRGFYRAY 420
QY 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449
DB 421 QAHVSSSNMAEFHASTWHTELTKTTDR 449

RESULT 7
AAB12579
ID AAB12579 standard. Protein: 449 AA.
XX
AC AAB12579;
XX
DT 09-NOV-2000 (first entry)
XX
DE Naphthalene dioxygenase mutant F352A protein sequence SEQ ID NO:33.
XX
KW Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KM inducible multi-component enzyme; alpha subunit; mutant; chiral diol;

```

XX	polymer;resin; pharmaceutical; rubber industry; bioremediation.
OS	Pseudomonas sp.
XX	Synthetic.
PN	WO200037480-A1.
PD	29-JUN-2000.
XX	
PF	26-OCT-1999; 99WO-US25079.
PR	26-OCT-1998; 98US-0105575.
XX	
PA	(IOWA ) UNIV IOWA RES FOUND.
XX	
PI	Parales R, Gibson D, Resnick S, Lee K;
DR	WPI: 2000-452174/39.
XX	
PT	N-PSDB; AAA65353.
XX	
PT	Novel naphthalene dioxygenase mutant having a specific amino acid
XX	substitution for preparing chiral diols for use in the polymer, resin,
XX	pharmaceutical or rubber industry and for carrying out bioremediation
XX	
PS	Claim 9; Page 128-129; 151pp; English.
XX	
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO
CC	related complex (I) comprising several polypeptides which contain an
CC	alpha subunit that contains substituted amino acids at specific
CC	positions. The polypeptides and host cells are useful for preparing
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-
CC	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-
CC	dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which
CC	involves contacting them with phenanthrene. The polypeptides and the
CC	host cells are also used in bioremediation in which they oxidise an
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,
CC	benzocyclopent-1-ene, anthracene, phenanthrene, dibenzof(1,4)dioxan,
CC	acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,
CC	dibenzothienophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a
CC	corresponding dihydro dihydroxy compound. The polypeptides and the host
CC	cells are useful for preparing chiral diols for use in the polymer,
CC	resin, pharmaceutical or rubber industry. The present sequence represents
CC	a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the
CC	exemplification of the present invention.
XX	
SO	Sequence 449 AA;
XX	
Query Match	99.8%; Score 2403; DB 21; Length 449;
Best Local Similarity	99.8%; Pred. No. 3.6e-224;
Matches 448; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
1	NNYNNKIIVSESGLSQKHLIHGDEDELFOHELTIIFARNMLFTHDSLIAPADGYTAKMG 60
Db	1 NMYNNKIIVSESGLSQKHLIHGDEDELFOHELTIIFARNMLFTHDSLIAPADGYTAKMG 60
QY	61 IDEVIVSRONDGSIRAFILVNCVCRHGRKTVLSVSEAGNAKGFVCSYHGMFGSGNELSVPE 120
Db	61 IDEVIVSRONDGSIRAFILVNCVCRHGRKTVLSVSEAGNAKGFVCSYHGMFGSGNELSVPE 120
QY	121 KOLYGESLKKCLGLKEVARVESFHGFTYGCDFDQEARPLMDYLGAAMYLEPMFKHSGGL 180
Db	121 KOLYGESLKKCLGLKEVARVESFHGFTYGCDFDQEARPLMDYLGAAMYLEPMFKHSGGL 180
QY	181 ELVGPGRKVIIVANKKAPARENFGADVAHYHGMTHASSIRGSEIFSSLAGNAALPREGACL 240
Db	181 ELVGPGRKVIIVANKKAPARENFGADVAHYHGMTHASSIRGSEIFSSLAGNAALPREGACL 240
QY	241 QMTSKYSGSGVGLMDQYSGVSHADVLPELIMAFGAKQERLNKEIGDVRARIYRSHLNTCTV 300

Db	241	QNTSKYGGGMVLWDGYSGVSHADLPEVLMAGGAQERLNEIGDVARIRYRSHLNCVY	300
QY	301	FPNNMMLTCSGVFVWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTIGPAGFWES	360
Db	301	FPNNSMLTCSGVFVWNPIDANTTEWYTAIYEKMPEDLKRRLADSVORTIGPAGFWES	360
QY	361	DDNDNMETASONGKKYOSRSDLLSNLGFGEVYGDVAVPGVYKSAIGETSYRGFYRAY	420
Db	361	DDNDNMETASONGKKYOSRSDLLSNLGFGEVYGDVAVPGVYKSAIGETSYRGFYRAY	420
QY	421	QAHVSSSNMAEFEEHASTHWTETLTCTDR	449
Db	421	QAHVSSSNMAEFEEHASTHWTETLTCTDR	449
RESULT 8			
AA12580			
ID	AA12580	standard; Protein; 449 AA.	
XX			
AC	AA12580;		
XX			
DT	09-NOV-2000	(first entry)	
XX			
DE	Naphthalene dioxygenase mutant F352T protein sequence SEQ ID NO:34.		
XX			
KM	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;		
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
XX	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
XX			
OS	Pseudomonas sp.		
XX	Synthetic.		
PN	MO200037480-A1.		
XX			
PD	29-JUN-2000.		
PF	26-OCT-1999;	99WO-US25079.	
XX			
PR	26-OCT-1998;	98US-0105575.	
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
PI	Parales R, Gibson D, Resnick S, Lee K;		
XX			
DR	WPI: 2000-452174/39.		
DR	N-PSDB: AAA65354.		
PT	Novel naphthalene dioxygenase mutant having a specific amino acid		
PT	substitution for preparing chiral diols for use in the polymer, resin,		
PT	pharmaceutical or rubber industry and for carrying out bioremediation		
PT	-		
PS	Claim 9; Page 129-131; 151pp; English.		
XX			
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO		
CC	related complex (I) comprising several polypeptides which contain an		
CC	alpha subunit that contains substituted amino acids at specific		
CC	positions. The polypeptides and host cells are useful for preparing		
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-diphenyl-3,4-		
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-		
CC	dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which		
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,		
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and		
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-		
CC	dihydronaphenanthrene or 3,4-dihydroxy-3,4-dihydronaphenanthrene which		
CC	involves contacting them with phenanthrene. The polypeptides and the		
CC	host cells are also used in bioremediation in which they oxidise an		
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,		
CC	benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,		
CC	acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,		
CC	dibenzofluorene, 9,10-dihydroanthracene, or 9,10-dihydronaphenanthrene a		
CC	corresponding dihydro dihydroxy compound. The polypeptides and the host		



CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.8%; Score 2403; DB 21; Length 449;  
 Best Local Similarity 99.8%; Pred. No. 3.6e-224;  
 Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MNYNKKILVSESGSKHLIHGDELFQHELKTIFARNMLFLTHDSLIPAPGDYVYAKMG 60
DB 1 MNYNKKILVSESGSKHLIHGDELFQHELKTIFARNMLFLTHDSLIPAPGDYVYAKMG 60
QY 61 IDEVIVSRONDGSIKRAFLVNCRRHGTILVSEGNKAGFVCSYHGWGFSNGELQSVPE 120
DB 61 IDEVIVSRONDGSIKRAFLVNCRRHGTILVSEGNKAGFVCSYHGWGFSNGELQSVPE 120
QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQEARPLMDYLGDAWYLEPMFKHSGGL 180
DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQEARPLMDYLGDAWYLEPMFKHSGGL 180
QY 181 ELVGPGRKVIYKANKKAPENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240
DB 181 ELVGPGRKVIYKANKKAPENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240
QY 241 QMTSKYSGMGVLMDSYGVHSADLPELMARFGAKOEKLNKEIGVRRIRYRSHLNCYV 300
DB 241 QMTSKYSGMGVLMDSYGVHSADLPELMARFGAKOEKLNKEIGVRRIRYRSHLNCYV 300
QY 301 FPNNSMLTCSGVYKWNPIDANTTEVMTYAIYEKMPEDLKRRLADSVORTIGPAGFWES 360
DB 301 FPNNSMLTCSGVYKWNPIDANTTEVMTYAIYEKMPEDLKRRLADSVORTIGPAGFWES 360
QY 361 DDNDNNETASQNGKKTQSQSDLSLNLGFGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420
DB 361 DDNDNNETASQNGKKTQSQSDLSLNLGFGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNNAEFHASTWHTELTKTTDR 449
DB 421 QAHVSSNNAEFHASTWHTELTKTTDR 449

RESULT 9
AAB12568
ID AAB12568 standard; Protein; 449 AA.
AC AAB12568;
XX
XX
XX 09-NOV-2000 (first entry)
DE NDO related complex alpha subunit protein sequence SEQ ID NO:16.
XX
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;
KW inducible multi-component enzyme; alpha subunit; mutant; chiral diol;
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.
XX
OS Pseudomonas sp.
OS Synthetic.
XX
XX WO200037480-A1.
XX
XX 29-JUN-2000.
XX
XX 26-OCT-1999; 99WO-US25079.
XX
XX 26-OCT-1998; 98US-0105575.
XX
XX (IOWA ) UNIV IOWA RES FOUNO.
XX
XX Parales R, Gibson D, Resnick S, Lee K.
XX

```

DR WPI; 2000-452174/39.

DR N-PSDB; AAA65342.

XX Novel naphthalene dioxygenase mutant having a specific amino acid  
 PT substitution for preparing chiral diols for use in the polymer, resin,  
 PT pharmaceutical or rubber industry and for carrying out bioremediation  
 PT

PS Claim 13; Page 102-103; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO  
 CC related complex (1) comprising several polypeptides which contain an  
 CC alpha subunit that contains substituted amino acids at specific  
 CC positions. The polypeptides and host cells are useful for preparing  
 CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
 CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
 CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
 CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
 CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
 CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
 CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
 CC involves contacting them with phenanthrene. The polypeptides and the  
 CC host cells are also used in bioremediation in which they oxidise an  
 CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
 CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzol(1,4)dioxan,  
 CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
 CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
 CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
 CC cells are useful for preparing chiral diols for use in the polymer,  
 CC resin, pharmaceutical or rubber industry. The present sequence represents  
 CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
 CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match 99.8%; Score 2402; DB 21; Length 449;  
 Best Local Similarity 99.6%; Pred. No. 4.5e-224;  
 Matches 447; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MNYNKKILVSESGSKHLIHGDELFQHELKTIFARNMLFLTHDSLIPAPGDYVYAKMG 60
DB 1 MNYNKKILVSESGSKHLIHGDELFQHELKTIFARNMLFLTHDSLIPAPGDYVYAKMG 60
QY 61 IDEVIVSRONDGSIKRAFLVNCRRHGTILVSEGNKAGFVCSYHGWGFSNGELQSVPE 120
DB 61 IDEVIVSRONDGSIKRAFLVNCRRHGTILVSEGNKAGFVCSYHGWGFSNGELQSVPE 120
QY 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQEARPLMDYLGDAWYLEPMFKHSGGL 180
DB 121 KDLYGSLNKKCLGLKEVARVESFHGFIYGCDFQEARPLMDYLGDAWYLEPMFKHSGGL 180
QY 181 ELVGPGRKVIYKANKKAPENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240
DB 181 ELVGPGRKVIYKANKKAPENFVGDAYHVGWTHASSLRSGESIFSSLAGNAALPPGAGL 240
QY 241 QMTSKYSGMGVLMDSYGVHSADLPELMARFGAKOEKLNKEIGVRRIRYRSHLNCYV 300
DB 241 QMTSKYSGMGVLMDSYGVHSADLPELMARFGAKOEKLNKEIGVRRIRYRSHLNCYV 300
QY 301 FPNNSMLTCSGVYKWNPIDANTTEVMTYAIYEKMPEDLKRRLADSVORTIGPAGFWES 360
DB 301 FPNNSMLTCSGVYKWNPIDANTTEVMTYAIYEKMPEDLKRRLADSVORTIGPAGFWES 360
QY 361 DDNDNNETASQNGKKTQSQSDLSLNLGFGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420
DB 361 DDNDNNETASQNGKKTQSQSDLSLNLGFGEDYVGDVAVPGVYKSAIGETSYRGFYRAY 420
QY 421 QAHVSSNNAEFHASTWHTELTKTTDR 449
DB 421 QAHVSSNNAEFHASTWHTELTKTTDR 449

```

RESULT 10

AB12583  
ID AAB12583 standard; Protein: 449 AA.  
XX  
AC AAB12583;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352W protein sequence SEQ ID NO:58.  
XX  
KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
PI WPI: 2000-452174/39.  
XX  
DR N-PSDB; AAA65376.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Example 7; Page 142-144; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX  
SQ Sequence 449 AA;  
Query Match 99.7%; Score 2401; DB 21; Length 449;  
Best Local Similarity 99.8%; Pred. No. 5.6e-224;  
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 61 IDEVTSRQNDGSTRAPLNVCRHGRKTLVSVEAGNAKGFVCSYHGMGFSGNGELQSVPE 120  
QY 121 KDLYGESLNKKKCLGLKEVARVESFHGFIYGCFFDQDAPPLMDYLGDAWYLEPMFKHSGGL 180  
Db 121 KDLYGESLNKKKCLGLKEVARVESFHGFIYGCFFDQDAPPLMDYLGDAWYLEPMFKHSGGL 180  
QY 181 ELVGPBGKVIKANKKAPAEENFVGDAYHVGMTTHASSLSRSGESIFSSLAGNAALPPEGAGL 240  
Db 181 ELVGPBGKVIKANKKAPAEENFVGDAYHVGMTTHASSLSRSGESIFSSLAGNAALPPEGAGL 240  
QY 241 QMTSKYSGMGVLMNGYSGVISADLVPELMAFGAKQERLNKEIGDVRARIYRSHLNCYV 300  
Db 241 QMTSKYSGMGVLMNGYSGVISADLVPELMAFGAKQERLNKEIGDVRARIYRSHLNCYV 300  
QY 301 FPNNSMLTFCSGFYKWNPIDANTTEVWTYAIYEKDMPEDLKRRLADSVQRTIGPAGFWS 360  
Db 301 FPNNSMLTFCSGFYKWNPIDANTTEVWTYAIYEKDMPEDLKRRLADSVQRTIGPAGFWS 360  
QY 361 DDNDMETSASONGKRRYQSRSDLSNLGFGEDVYGDAYVPGVKSAGTSTYRGFYRAY 420  
Db 361 DDNDMETSASONGKRRYQSRSDLSNLGFGEDVYGDAYVPGVKSAGTSTYRGFYRAY 420  
QY 421 QAHVSSNMWAEFEHASTWHTLTKTTDR 449  
Db 421 QAHVSSNMWAEFEHASTWHTLTKTTDR 449  
RESULT 11  
AAB12578  
ID AAB12578 standard; Protein: 449 AA.  
XX  
AC AAB12578;  
XX  
DT 09-NOV-2000 (first entry)  
XX  
DE Naphthalene dioxygenase mutant F352G protein sequence SEQ ID NO:32.  
XX  
KM Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
KM Inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KM polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
XX  
PN WO200037480-A1.  
XX  
PD 29-JUN-2000.  
XX  
PF 26-OCT-1999; 99WO-US25079.  
XX  
PR 26-OCT-1998; 98US-0105575.  
XX  
PA (IOWA ) UNIV IOWA RES FOUND.  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX  
PI WPI: 2000-452174/39.  
XX  
DR N-PSDB; AAA65352.  
XX  
PT Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
PS Claim 9; Page 126-127; 151pp; English.  
XX  
CC The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which

CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthanthrene or 3,4-dihydroxy-3,4-dihydronaphthanthrene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzof(1,4)dioxan,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthanthrene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.

XX Sequence 449 AA:

Query Match	99.7%	Score 2400;	DB 21;	Length 449;
Best Local Similarity	99.8%	Prod. No. 7e-224;		
Matches 448;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0

Qy	1	MNYNNKLIVYSSGSLXOKHLHJGDELEFÖHEKLTTPANNWLFJLTHDSJLTPADGYVYAKMG	60
Dd	1	MNYNNKLIVYSSGSLXÖHHLHJGDELEFÖHEKLTTPANNWLFJLTHDSJLTPADGYVYAKMG	60
Qy	61	IDEVYVSQNDGSIIRAFJLNCVRRHGKTLVSEAGNAKGFVCSYHGWFSGNGLQSVPE	120
Dd	61	IDEVYVSQNDGSIIRAFJLNCVRRHGKTLVSEAGNAKGFVCSYHGWFSGNGLQSVPE	120
Qy	121	KDLVYGSJLNKKCLGLKEVARYESFHGTYGCFDQAEAPJLMDYJGDAWYJLEPMFKHSGGJ	180
Dd	121	KDLVYGSJLNKKCLGLKEVARYESFHGTYGCFDQAEAPJLMDYJGDAWYJLEPMFKHSGGJ	180
Qy	181	ELVBPBGVYVYKANKKAPAEFNFYDAVHYCMTTHASSJLRGSEJFSSJLAGNAALPPRGAGJ	240
Dd	181	ELVBPBGVYVYKANKKAPAEFNFYDAVHYCMTTHASSJLRGSEJFSSJLAGNAALPPRGAGJ	240
Qy	241	QMTSKYSGSGMVLMDGYVSGVSHADLVPELMAFGAKQERLNKEIGDVRARIYRSHLNCYV	300
Dd	241	QMTSKYSGSGMVLMDGYVSGVSHADLVPELMAFGAKQERLNKEIGDVRARIYRSHLNCYV	300
Qy	301	FPNNSMLTCSGVFVWNPIDANTTEVWYTAJVEKDMPEDJLKRRLABSVQRTTGPAGFWES	360
Dd	301	FPNNSMLTCSGVFVWNPIDANTTEVWYTAJVEKDMPEDJLKRRLABSVQRTTGPAGFWES	360
Qy	361	DDNNMMEYASQNGKRYOSRPSDLSJNLGFEDYVGDVAVGYGVKSAJGSETYRGFYRAY	420
Dd	361	DDNNMMEYASQNGKRYOSRPSDLSJNLGFEDYVGDVAVGYGVKSAJGSETYRGFYRAY	420
Qy	421	QAHVSSSNMAEEFHASSTWHTLTKTDDR 449	
Dd	421	QAHVSSSNMAEEFHASSTWHTLTKTDDR 449	

Accession	Protein	Gene	Organism	Source	Accession	Protein	Gene	Organism	Source
AB012584	standard; Protein; 448 AA.				AB012584	standard; Protein; 448 AA.			
AB012584					AB012584				
09-NOV-2000	(first entry)				09-NOV-2000	(first entry)			
Naphthalene dioxygenase mutant F352Y	protein sequence SEQ ID NO:59.				Naphthalene dioxygenase mutant F352Y	protein sequence SEQ ID NO:59.			
Pseudomonas sp. strain NCIB 9816-4	naphthalene dioxygenase; NDO;				Pseudomonas sp. strain NCIB 9816-4	naphthalene dioxygenase; NDO;			
inducible multi-component enzyme; alpha subunit; mutant; chiral diol;					inducible multi-component enzyme; alpha subunit; mutant; chiral diol;				
Polymer; resin; pharmaceutical; rubber industry; bioremediation.					Polymer; resin; pharmaceutical; rubber industry; bioremediation.				
Pseudomonas sp.					Pseudomonas sp.				
Synthetic.					Synthetic.				
MO200037480-A1.					MO200037480-A1.				

XX 29-JUN-2000.  
PD  
XX 26-OCT-1999; 99MO-US25079.  
PF  
XX 26-OCT-1998; 98US-0105575.  
PR  
XX (IOWA ) UNIV IOWA RES FOUND.  
PA  
XX  
XX  
PI Parales R, Gibson D, Resnick S, Lee K;  
XX WPI: 2000-452174/39.  
DR N-PSDB: AAA65377.  
DR  
XX  
XX Novel naphthalene dioxygenase mutant having a specific amino acid  
PT substitution for preparing chiral diols for use in the polymer, resin,  
PT pharmaceutical or rubber industry and for carrying out bioremediation  
PT  
XX  
XX Example 7, Page 144-145; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO related complex (I) comprising several polypeptides which contain an alpha subunit that contains substituted amino acids at specific positions. The polypeptides and host cells are useful for preparing (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-dihydroxyindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which involves contacting them with naphthalene, biphenyl, phenanthrene, indene, and 1,2-dihydronaphthalene respectively. The polypeptides and the host cells are also useful for preparing 1,2-dihydroxy-1,2-dihydronaphanthrene or 3,4-dihydroxy-3,4-dihydronaphanthrene which involves contacting them with phenanthrene. The polypeptides and the host cells are also used in bioremediation in which they oxidise an aromatic compound such as indene, 1,2-dihydronaphthalene, benzo(c)chlept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan, acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran, dibenzochlorophene, 9,10-dihydroanthracene, or 9,10-dihydronaphanthrene a corresponding dihydro dihydroxy compound. The polypeptides and the host cells are useful for preparing chiral diols for use in the polymer, resin, pharmaceutical or rubber industry. The present sequence represents a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the exemplification of the present invention.

Qy	1	MNNKTLTVSSGSLQKHLHGDDEEHOHEKLTFAANNLFLTHDSILIPAGDYVAAKMG	60
Dp	1	MNNKTLTVSSGSLQKHLHGDDEEHOHEKLTFAANNLFLTHDSILIPAGDYVAAKMG	60
Qy	61	IDEVIVSRQNDGSLRAFLNVCRRHGKTLVSYEAGNAGFCVSYHGWFSGNGELQSVPE	120
Dp	61	IDEVIVSRQNDGSLRAFLNVCRRHGKTLVSYEAGNAGFCVSYHGWFSGNGELQSVPE	120
Qy	121	KDLVGSSTNKKCLGLKEVARYESHGTYGCFDEAPRLMDYLGDAWYLEPMPKHSGL	180
Dp	121	KDLVGSSTNKKCLGLKEVARYESHGTYGCFDEAPRLMDYLGDAWYLEPMPKHSGL	180
Qy	181	ELVGPBGKVVIAKAMWKAPAEFNVDAADHVGVTHASSLSRGSSTFSSLAGNAALPPEGAGL	240
Dp	181	ELVGPBGKVVIAKAMWKAPAEFNVDAADHVGVTHASSLSRGSSTFSSLAGNAALPPEGAGL	240
Qy	241	QMTSKYSGMCLVDGYSGVNSADLVPELMAFGGAKOERLNKETGDVRIATYESHNLCTV	300
Dp	241	QMTSKYSGMCLVDGYSGVNSADLVPELMAFGGAKOERLNKETGDVRIATYESHNLCTV	300
Qy	301	FPNNSMLTLCSCGVFYVWMPIDANTLEVTVYALVEKDMPEDLKRRLADSVORTIPAGFWES	360
Dp	301	FPNNSMLTLCSCGVFYVWMPIDANTLEVTVYALVEKDMPEDLKRRLADSVORTIPAGFWES	360

Oy	361	DNDNNMTASQNGKRRQSRDLSLNLGFEEDYDGAIVYGVGKSAIGETSYRGFFRAY	420
Db	361	DNDNNMTASQNGKRRQSRDLSLNLGFEEDYDGAIVYGVGKSAIGETSYRGFFRAY	420
Oy	421	QAHVSSNNAAEFHASTWHTLTKTTD	448
Db	421	QAHVSSNNAAEFHASTWHTLTKTTD	448
RESULT 13			
ID	AAB12569		
	AAB12569	strand: protein; 449 AA.	
XX			
AC	AAB12569;		
DT	09-NOV-2000	(first entry)	
DE	NDO related complex alpha subunit protein sequence SEQ ID NO:17.		
KM	Pseudomonas sp. strain NCIB 9816-4, naphthalene dioxygenase: NDO;		
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;		
XX	polymer; resin; pharmaceutical; rubber industry; bioremediation.		
OS	Pseudomonas sp.		
XX	Synthetic.		
PN	WO200037480-A1.		
PD	29-JUN-2000.		
XX			
PF	26-OCT-1999;	99WO-US25079.	
XX			
PR	26-OCT-1998;	98US-0105575.	
XX			
PA	(IOWA ) UNIV IOWA RES FOUND.		
PI	Parales R, Gibson D, Resnick S, Lee K;		
DR	WPI: 2000-452174/39.		
XX	N-PSDB: AAA65343.		
PT	Novel naphthalene dioxygenase mutant having a specific amino acid		
PT	substitution for preparing chiral diols for use in the polymer, resin,		
PT	pharmaceutical or rubber industry and for carrying out bioremediation		
XX	-		
PS	Claim 13; Page 104-105; 151pp; English.		
XX			
CC	The present invention describes a naphthalene dioxygenase (NDO) or NDO		
CC	related complex (I) comprising several polypeptides which contain an		
CC	alpha subunit that contains substituted amino acids at specific		
CC	positions. The polypeptides and host cells are useful for preparing		
CC	(-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-diphenyl-3,4-		
CC	dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-		
CC	dihydroxylindan, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which		
CC	involves contacting them with naphthalene, biphenyl, phenanthrene,		
CC	indene, and 1,2-dihydronaphthalene respectively. The polypeptides and		
CC	the host cells are also useful for preparing 1,2-dihydroxy-1,2-		
CC	dihydronaphthanrene or 3,4-dihydroxy-3,4-dihydronaphthanrene which		
CC	involves contacting them with phenanthrene. The polypeptides and the		
CC	host cells are also used in bioremediation in which they oxidise an		
CC	aromatic compound such as indene, 1,2-dihydronaphthalene,		
CC	benzocyclohept-1-ene, anthracene, phenanthrene, dibenzo(1,4)dioxan,		
CC	acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,		
CC	dibenzofluorene, 9,10-dihydroanthracene, or 9,10-dihydronaphthanrene a		
CC	corresponding dihydro dihydroxy compound. The polypeptides and the host		
CC	cells are useful for preparing chiral diols for use in the polymer,		
CC	resin, pharmaceutical or rubber industry. The present sequence represents		
CC	a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the		
CC	exemplification of the present invention.		
XX			
Sequence	449 AA;		

	Query Match	99.5%;	Score 2397;	DB: 21;	Length 449;
	Best Local Similarity	99.38;	Pred. No. 1,4e-223;		
	Matches 446;	Conservative 2;	Mismatches 1;	Indels 0;	Gaps
OY	1 MMYNNKILVSESGLSQKHILHIGDEELFOHELKTTFARNWLEFLTHDSLIPAPGDYVAKMG	60			
Db	1 MMYNNKILVSESGLSQKHILHIGDEELFOHELKTTFARNWLEFLTHDSLIPAPGDYVAKMG	60			
OY	61 IDEIVTSRNDGSIIRAFPLNVCRRHGKTLVSVEAGNAKGFWCSYHGMFGSNGELOSVPPE	120			
Db	61 IDEIVTSRNDGSIIRAFPLNVCRRHGKTLVSVEAGNAKGFWCSYHGMFGSNGELOSVPPE	120			
OY	121 KDLVESSLNKCKCLGKEVARVESFHGFYTCGCFDDEAPRLMDYLGDAWYLEPMFKHSGGL	180			
Db	121 KDLVESSLNKCKCLGKEVARVESFHGFYTCGCFDDEAPRLMDYLGDAWYLEPMFKHSGGL	180			
OY	181 ELVGPGKVVIKANKKAFAENFVDADYHVGTHTASSLRGSESIFSSLAGMAALPPREGAQL	240			
Db	181 ELVGPGKVVIKANKKAFAENFVDADYHVGTHTASSLRGSESIFSSLAGMAALPPREGAQL	240			
OY	241 QMTSKYSGGMGYLWDGYSGVSHADLVPELAAFGAKOERLNKEIGVRRARIYRSHLNTCV	300			
Db	241 QMTSKYSGGMGYLWDGYSGVSHADLVPELAAFGAKOERLNKEIGVRRARIYRSHLNTCV	300			
OY	301 PNNNSMLTFCSGFEKYWNPIDANTTEVMVTYAIVEKDMPEDLKRLADSYORTIGPAGFMES	360			
Db	301 PNNNSMLTFCSGFEKYWNPIDANTTEVMVTYAIVEKDMPEDLKRLADSYORTIGPAGFMES	360			
OY	361 DDNDMMETASONGKKYQSQRSDLLSNLGFGEDEVGDAYVPGVYGKSALIGETSYRGFYRAY	420			
Db	361 DDNDMMETASONGKKYQSQRSDLLSNLGFGEDEVGDAYVPGVYGKSALIGETSYRGFYRAY	420			
OY	421 OAHVSSSNMAREFEHASSTWHTELTKTTDR	449			
Db	421 RAHVSSSNMAREFEHASSTWHTELTKTTDR	449			
<b>RESULT 14</b>					
AAB12570	standard; Protein: 449 AA.				
XX					
AC	AAB12570;				
XX					
DT	09-NOV-2000 (first entry)				
XX					
DE	NDO related complex alpha subunit protein sequence SEQ ID NO:18.				
XX					
KW	Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;				
KW	inducible multi-component enzyme; alpha subunit; mutant; chiral diol;				
XX	polymer; resin; pharmaceutical; rubber industry; bioremediation.				
OS	Pseudomonas sp.				
OS	Synthetic.				
PN	WO200037480-A1.				
XX					
PD	29-JUN-2000.				
XX					
PF	26-OCT-1999; 99WO-US25079.				
XX					
PR	26-OCT-1998; 98US-0105575.				
XX					
PA	(IOWA ) UNIV IOWA RES FOUND.				
XX					
PI	Parales R, Gibson D, Resnick S, Lee K;				
XX					
XX	WPI; 2000-452174/39.				
DR	N-PSTD; AAA65344.				
PT	Novel naphthalene dioxygenase mutant having a specific amino acid				
PT	substitution for preparing chiral diols for use in the polymer, resin,				
PT	pharmaceutical or rubber industry and for carrying out bioremediation				

PT -  
XX Claim 13; Page 105-107; 151pp; English.  
PS  
XX The present invention describes a naphthalene dioxygenase (NDO) or NDO  
XX related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzofuran,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX

Seq Sequence 449 AA:

Query Match 97.5%; Score 2349; DB 21; Length 449;  
Best Local Similarity 96.7%; Pred. No. 6.3e-219;  
Matches 434; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNYNKKILVSESGLSQKHLIHGDEELFQHELTIFARNMLFLTHDSLTPAPDYYTAKNG 60  
DB 1 MNYNKKILVSESGLSQKHLIHGDEELFQHELTIFARNMLFLTHDSLTPAPDYYTAKNG 60  
QY 61 IDEVIVSRQNDISIRAFNLVNCRRHCKTLVSVAGNAKGFVCSYHGMGFSGNGLSVPE 120  
DB 61 IDEVIVSRQNDISIRAFNLVNCRRHCKTLVSVAGNAKGFVCSYHGMGFSGNGLSVPE 120  
QY 121 KDLYESLNKKKCLGLKEVARVESFHGFTYGCDFQDAPPLMDYLDGAAMYLEPMFKHSGGL 180  
DB 121 KDLYESLNKKKCLGLKEVARVESFHGFTYGCDFQDAPPLMDYLDGAAMYLEPMFKHSGGL 180  
QY 121 KELYESLNKKKCLGLKEVARVESFHGFTYGCDFQDAPPLMDYLDGAAMYLEPMFKHSGGL 180  
DB 121 KELYESLNKKKCLGLKEVARVESFHGFTYGCDFQDAPPLMDYLDGAAMYLEPMFKHSGGL 180  
QY 181 ELVGPGRKVIKANKKAPENFVGDAYHVGWTHASSLSRGSSEIFASLAGNAVLPEPGAGL 240  
DB 181 ELVGPGRKVIKANKKAPENFVGDAYHVGWTHASSLSRGSSEIFASLAGNAVLPEPGAGL 240  
QY 241 QMTSKYSGMGVLMQDYGSGVSHSADLVPELMAFGAKOERLNEIGDVRARIYRSHLNCYV 300  
DB 241 QMTSKYSGMGVLMQDYGSGVSHSADLVPELMAFGAKOERLNEIGDVRARIYRSHLNCYV 300  
QY 301 PPNNSMLTCSGVFKVWNPIDANTEWTYAIYEKDPEDLKRRLADSVORTTGPAGFWES 360  
DB 301 PPNNSMLTCSGVFKVWNPIDANTEWTYAIYEKDPEDLKRRLADSVORTTGPAGFWES 360  
QY 361 DDNDMMETASONGKRYQSDLSNLGEGEDYVGDAYPGVGSALGETSYRGRFYRAY 420  
DB 361 DDNDMMETASONGKRYQSDLSNLGEGEDYVGDAYPGVGSALGETSYRGRFYRAY 420  
QY 421 QAHVSSSNMAEFERHASTWHELTJKTTR 449  
DB 421 QAHVSSSNMAEFERHASTWHELTJKTTR 449

RESULT 15  
AAB12571  
ID AAB12571 standard; Protein; 449 AA.  
XX AAB12571;  
XX  
DT 09-NOV-2000 (first entry)

XX NDO related complex alpha subunit protein sequence SEQ ID NO:19.  
DE  
XX Pseudomonas sp. strain NCIB 9816-4; naphthalene dioxygenase; NDO;  
XX inducible multi-component enzyme; alpha subunit; mutant; chiral diol;  
KW polymer; resin; pharmaceutical; rubber industry; bioremediation.  
XX  
OS Pseudomonas sp.  
OS Synthetic.  
FH  
FT Key Location/Qualifiers  
FT Misc-difference 35 /note="unspecified"  
PD WO200037480-A1.  
PD 29-JUN-2000.  
XX  
XX 26-OCT-1999; 99WO-US25079.  
XX  
XX 26-OCT-1998; 98US-0105575.  
XX  
XX (IOWA ) UNIV IOWA RES FOUND.  
XX  
XX Parales R, Gibson D, Resnick S, Lee K;  
XX  
XX WPI: 2000-452174/39.  
XX  
XX N-PSDB; AAA63545.  
XX  
XX Novel naphthalene dioxygenase mutant having a specific amino acid  
XX substitution for preparing chiral diols for use in the polymer, resin,  
XX pharmaceutical or rubber industry and for carrying out bioremediation  
XX  
XX

Claim 13; Page 107-109; 151pp; English.

The present invention describes a naphthalene dioxygenase (NDO) or NDO  
CC related complex (I) comprising several polypeptides which contain an  
CC alpha subunit that contains substituted amino acids at specific  
CC positions. The polypeptides and host cells are useful for preparing  
CC (-)-(1S,2R)-cis-naphthalene dihydrodiol, (- or +)-cis-biphenyl-3,4-  
CC dihydrodiol, (1S,2R)-cis-phenanthrene-1,2-dihydrodiol, cis-1,2-  
CC dihydroxyindane, 1,2-dihydroxy-1,2,3,4-tetrahydronaphthalene, which  
CC involves contacting them with naphthalene, biphenyl, phenanthrene,  
CC indene, and 1,2-dihydronaphthalene respectively. The polypeptides and  
CC the host cells are also useful for preparing 1,2-dihydroxy-1,2-  
CC dihydronaphthalene or 3,4-dihydroxy-3,4-dihydronaphthalene which  
CC involves contacting them with phenanthrene. The polypeptides and the  
CC host cells are also used in bioremediation in which they oxidise an  
CC aromatic compound such as indene, 1,2-dihydronaphthalene,  
CC benzocyclohept-1-ene, anthracene, phenanthrene, dibenzofuran,  
CC acenaphthylene, naphthalene, biphenyl, fluorene, dibenzofuran,  
CC dibenzothiophene, 9,10-dihydroanthracene, or 9,10-dihydronaphthalene a  
CC corresponding dihydro dihydroxy compound. The polypeptides and the host  
CC cells are useful for preparing chiral diols for use in the polymer,  
CC resin, pharmaceutical or rubber industry. The present sequence represents  
CC a naphthalene dioxygenase (NDO) mutant alpha subunit, given in the  
CC exemplification of the present invention.  
XX  
XX

Seq Sequence 449 AA:  
Query Match 96.4%; Score 2321; DB 21; Length 449;  
Best Local Similarity 95.5%; Pred. No. 3.3e-216;  
Matches 429; Conservative 13; Mismatches 7; Indels 0; Gaps 0;

QY 1 MNYNKKILVSESGLSQKHLIHGDEELFQHELTIFARNMLFLTHDSLTPAPDYYTAKNG 60  
DB 1 MNYNKKILVSESGLSQKHLIHGDEELFQHELTIFARNMLFLTHDSLTPAPDYYTAKNG 60  
QY 61 IDEVIVSRQNDISIRAFNLVNCRRHCKTLVSVAGNAKGFVCSYHGMGFSGNGLSVPE 120  
DB 61 IDEVIVSRQNDISIRAFNLVNCRRHCKTLVSVAGNAKGFVCSYHGMGFSGNGLSVPE 120

Oy	121	KDLVGESENKKCLGIEKVARARESHGFLYCGFDEAPPLIDGADAMYLEPMEKHSGL	180
Dp	121	KELVGESENKKCLGIEKVARARESHGFLYCGFDEAPSLINDYLCGDAMYLEPFLFKHSGL	180
Oy	181	ELVGPBGKVVTKANMKAPAEFNEVDAYHVGMTTHASSLSRSGESIFSSLAGNAALPPEGAGL	240
Dp	181	ELVGPBGKVVTKANMKAPAEFNEVDAYHVGMTTHASSLSRSGESIFSSLAGNAALPPEGAGL	240
Oy	241	QMTSKYSGSGMGLMDGYSVGSVHADLVPBELMAFGGAKOERLNKEIGDYRARIRYSHLNCIV	300
Dp	241	QMTSKYSGSGMGLMDGYSVGSVHADLVPBELMAFGGAKOERLNKEIGDYRARIRYSHLNCIV	300
Oy	301	FPNNSMLTCSGVFKYKWNPIDANTEVWTYALAEKMDPEDLKRLRADSVOYRTTIPAGFWES	360
Dp	301	FPNNSVLTCSGVFKYKWNPIDANTEVWTYALAEKMDPEDLKRLRADSVOYRTTIPAGFWES	360
Oy	361	DONDMETASONGKKYKOSRSDLSLNLGFGGDYGDAAVYQGVKKSALIGTSVRYGFYRAY	420
Dp	361	DONDMETASONGKKYKOSRSDLSLNLGFGGDYGDAAVYQGVKKSALIGTSVRYGFYRAY	420
Oy	421	CAHVSSSNMAEFKSHASTWHTLELKTDDR	449
Dp	421	CAHVSSSNMAEFKSHASTWHTLELKTDDR	449

Search completed: January 27, 2003, 08:58:53  
Job time : 31.1667 secs